Contract No.: 500-94-0062
MPR Reference No.: 8246-123

MATHEMATICA
Policy Research, Inc.

The Impacts of Per-Episode Prospective Payment for Medicare Home Health Care on the Quality of Care: Less Is Not Necessarily Worse

May 9, 2000

Arnold Chen

Programmers:

Helaine Noveck
Elizabeth Stuart
Will Black
Amy Zambrowski

Survey Director:
Rhoda Cohen

Submitted to:

Health Care Financing Administration
7500 Security Blvd, (C-3-21-06)
Baltimore, MD 21244-1850

Project Officer:
Ann Meadow

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
(609) 799-3535

Project Director:
Valerie Cheh

Principal Investigators:
Randall Brown
Valerie Cheh
Barbara Phillips

ACKNOWLEDGMENTS

I am grateful for help from many colleagues in preparing this report. Valerie Cheh, Barbara Phillips, and Randall Brown carefully read earlier drafts and made insightful and critical comments. Randall Brown patiently provided guidance in the analysis and statistical methodology. Chris Trenholm also made valuable suggestions for the analysis. Barbara Phillips, Jennifer Schore, and Chris Trenholm contributed to the background chapters through previous work. Barbara Phillips authored Appendix A on the design, sampling, and fielding of the patient surveys. Helaine Noveck, Elizabeth Stuart, and Amy Zambrowski provided outstanding programming support. Laura Berenson meticulously and skillfully edited the report with help from Patricia Ciaccio. Jennifer Baskwell tirelessly typed, formatted, and produced the draft report with help from Jill Miller and Stephanie Collins, and Cindy McClure made the revisions to and produced the final report.

I also thank Ann Meadow, our project officer at HCFA for ongoing advice and assistance, and the Center for Health Policy Research at the University of Colorado for providing the Quality Assurance and documentation. Errors or omissions are, of course, my responsibility.

CONTENTS

CONTENTS *(continued)*

CONTENTS *(continued)*

CONTENTS *(continued)*

TABLES

TABLES *(continued)*

TABLES *(continued)*

EXECUTIVE SUMMARY

This report investigates whether the quality of home health care changed under the Per-Episode Home Health Prospective Payment Demonstration. The demonstration was part of ongoing efforts by the Health Care Financing Administration (HCFA) to improve the cost-effectiveness of the Medicare program. Under the demonstration, home health agencies in the treatment group were paid a fixed, lump-sum payment for the first 120 days of each episode of care they provided to Medicare beneficiaries, regardless of the number or cost of the visits. By permitting agencies to retain most of any surplus payments over their cost, this payment method gave agencies an incentive to provide home health care in a cost-efficient manner.

We detected no adverse impacts of prospective payment on numerous measures of patient health, function, or health services use, despite the substantial reduction in home health visits resulting from prospective payment. We found weak evidence suggesting that emergency visits to clinics and physicians' offices decreased as a result of prospective payment, and that beneficiaries were slightly more dissatisfied with the interpersonal care that agency staff provided. We found no evidence that prospective payment affected different types of agencies or different types of patients differentially.

BACKGROUND

Ninety-one agencies in five states voluntarily entered the three-year demonstration at the start of their 1996 fiscal year and were randomly assigned to either the demonstration's prospective payment method (the treatment group) or Medicare's usual method of cost-based reimbursement (the control group). The lump-sum payment amount that treatment agencies received for each home health admission was based on each agency's own costs for the fiscal year preceding its entry into the demonstration, adjusted for inflation, with an additional retrospective adjustment for case-mix change.[1] An agency could start a new episode for a previously admitted patient (and thus receive a new payment) only after the 120-day "at-risk" period had ended and a 45-day gap in services had taken place. Agencies were paid a fixed rate for visits made after the at-risk period (but which still were part of the same episode). The rate for these so-called "outlier" visits was based on the type of visit and the agency's predemonstration costs. HCFA reimbursed treatment agencies for up to 99 percent of fiscal year losses up to the Section 223 cost-per-visit limits for all Medicare home health agencies. Agencies had to share profits in excess of specified limits with HCFA.

[1]Some exceptions apply. For example, durable medical equipment, nonroutine medical supplies, and Part B ambulatory home health services continued to be reimbursed at cost.

RESEARCH QUESTIONS AND METHODOLOGY

We investigated whether prospective payment affected a broad spectrum of patient outcomes obtained from three data sources: (1) Medicare claims data, (2) demonstration quality assurance data, and (3) a telephone survey of home health patients. The five main types of outcomes were:

1. ***Self-Reported General Health Status.*** Whether perceived health was good or excellent, number of days restricted to bed, and satisfaction with life

2. ***Functional Status.*** Basic activities of daily living (grooming, bathing, toileting, transferring, and ambulating) and instrumental activities of daily living (preparing light meals, housekeeping, and managing oral medications)

3. ***Medical Symptoms and Outcomes.*** Pain, pressure ulcers, wound status, dyspnea, urinary incontinence, confusion, behavioral problems, and mortality

4. ***Use of Health Services.*** Emergency care, admission to hospital for a diagnosis related to the home health care diagnosis (same-body-system diagnosis), admission to skilled nursing facility for a same-body-system diagnosis, and readmission to home health care for a same-body-system diagnosis

5. ***Satisfaction with Agency Care.*** Overall satisfaction with care received and satisfaction with specific aspects of care

We used logit models to estimate effects of prospective payment on the outcome variables while controlling for preexisting differences between treatment and control agencies or their patients that may have remained despite randomization. We obtained the variables to control for preexisting treatment and control differences in patient characteristics from case-mix adjustment data and Medicare claims data. Agency characteristics were drawn from agency cost reports, the demonstration implementation contractor, and the Area Resource File. Observations were weighted so that each agency was represented equally in the analysis. When estimating the standard errors of the estimated treatment-control differences, we accounted for the effects of sample clustering and weighting. In sensitivity analyses, our impact estimates were robust to both an alternative weighting scheme and an alternative sample of patients.

FINDINGS

Prospective Payment Had No Effect on Multiple Measures of Health and Functional Status

We consistently found throughout our analysis that prospective payment had no discernable effects on the quality of care. We found no differences between patients of treatment agencies and patients of control agencies in their perceptions of overall health or in self-reports of the number of days confined to bed. We also found little evidence to support a finding of impacts of prospective

payment on either basic or instrumental activities of daily living. Moreover, we concluded that prospective payment had no impacts on a wide range of medical symptoms and outcomes (such as pain, dyspnea, and mortality). We also found no demonstration impacts on admission to skilled nursing facilities or home health agencies for same-body-system diagnoses.

We did find indications that prospective payment may have led to a reduction in some types of health service use. Roughly 1.25 percent of patients of cost-reimbursed agencies had an emergency visit to a hospital outpatient clinic, compared with roughly 0.75 percent of patients of prospectively paid agencies. Three percent of patients of cost-reimbursed agencies had an emergency visit to a physician's office, compared with two percent of patients of prospectively paid agencies. We observed no effect on visits to hospital emergency rooms. A reduction in hospitalizations for same-body-system diagnoses among the treatment group was evident by 120 days after home health admission. This reduction widened over time so that, by one year, 35 percent of patients of cost-reimbursed agencies had had a same-body-system hospitalization, compared with 33 percent of patients of prospectively paid agencies.

We are cautious about accepting these treatment-control differences in emergency and hospital care as real impacts. On the one hand, a previous report on this demonstration found no impact on either the use of most Medicare services or total Medicare costs (Schore 1999). Furthermore, the differences we observed may reflect only a baseline difference between prospectively paid and cost-reimbursed agencies in predemonstration rates of hospitalization, rather than an impact of prospective payment. On the other hand, additional analyses that statistically controlled for agency-level base-quarter hospital use did not completely eliminate the observed differences. One explanation for real differences is that treatment agencies did a better job of preventing medical problems from occurring and of educating their patients in self-care. However, treatment agency patients were slightly *more* dissatisfied with the amount of attention and encouragement they received from agency staff, and we did not observe any other indications that treatment agencies provided higher-quality care. A second explanation is that relatively less contact by treatment agency staff had the effect of reducing the number of unnecessary agency staff referrals for emergency room and hospital care. Both the lack of impacts on patient outcomes and findings from previous studies--that increased medical surveillance led to increased health service use--support this explanation. However, our main conclusion is that the reductions in home health visits arising from prospective payment did not cause emergency visits or hospital admissions to *increase*.

Prospective payment did have a small impact on patients' satisfaction with interpersonal aspects of care. Large majorities of treatment agency and control agency patients were equally satisfied with the overall care their agency provided. However, a small fraction of patients were dissatisfied with specific aspects of interpersonal care, and patients of prospectively paid agencies were somewhat more likely to be dissatisfied than were patients of control agencies. Six percent of treatment agency patients felt that staff rushed through their work, compared with roughly four percent of control agency patients. Eleven percent of treatment agency patients felt that staff did not encourage them to be independent, compared with roughly 8 percent of control agency patients. In addition, eight percent of treatment agency patients felt that staff did not pay attention to them, compared with

roughly five percent of control agency patients.[2] Although these differences are proportionally large and statistically significant, they represent so few patients that policymakers may consider the increase in dissatisfaction a small price to pay for the significant utilization decrease resulting from prospective payment

Prospective Payment Had No Effects on Agency or Patient Subgroups

Patient outcomes may have varied across agency or patient subgroups because the incentives of prospective payment caused certain types of agencies to behave differently or certain types of patients to be treated differently. For example, for-profit agencies may have responded more eagerly than nonprofit agencies to the profit incentives of the demonstration, or agencies may have reduced the number of visits by a greater amount for patients who had informal caregivers available to provide assistance than for patients who lacked such assistance. To investigate this possibility, we studied the impact of the demonstration on five types of agencies and on three patient subgroups. The five types of agencies studied were (1) agencies with high-use or low-use predemonstration practice patterns, (2) for profit or nonprofit, (3) small or large, (4) hospital based or freestanding (auspice), and (5) below or above cost limits. The three patient-level subgroups (1) were independent or not independent in taking oral medications, (2) had another caregiver available or did not, and (3) had high or low expected costs.

We found no credible evidence that prospective payment had a differential effect on any particular type of agency for any of the several outcomes studied (that is, health services use, health, or functional status). We observed scattered differences across subgroups that occasionally were statistically significant, but that varied in direction and showed no consistent pattern. These findings for patient outcomes contrasted to those for home health care service use, where we had found suggestions that prospective payment led to greater reductions in service use for certain agency subgroups (high use, small, or freestanding agencies, and agencies above cost limits).

Furthermore, we also found no evidence of any patient-level subgroup effects. We had previously found that the effects of prospective payment on home health care service use varied little with patient characteristics, and we likewise observed no subgroup impacts on the patient quality-of-care outcomes in any of the patient-level subgroups.

[2]The percentages for control group agencies are unadjusted control group means. We derived the percentages for treatment group agencies by adding the estimated treatment-control difference to the unadjusted control group means.

LIMITATIONS OF THE ANALYSIS

The Study May Not Have Detected Small Demonstration Effects

As with all evaluations, our study runs the risk of producing "false-negative" results (that is, the inability to detect true program impacts because of inadequate statistical power). However, for most of the outcomes studied, we had sufficient power (even after accounting for design effects) to detect treatment-control differences of four to five percentage points at 80 percent power. These minimum detectable differences are large enough to be of substantive or policy-relevant size.

It is possible that our study did not measure additional outcomes that would have been more sensitive to decreases in the number of home health visits. However, given the wide array of outcomes we did measure, this possibility is unlikely.

A National Payment System Might Differ from the Demonstration Payment System

A national program of prospective payment could lead to even larger service reductions than were observed in this demonstration, with unknown effects on the quality of care. For example, under a national system, agencies would not be protected from incurring financial losses, which might lead some of them to respond more aggressively to program incentives. Moreover, an agency's payment would almost certainly not be based on its own previous cost per episode, but rather, on national averages or on a blend of regional and national averages. To lower their costs per episode enough to remain solvent, many agencies, such as those with high-use practice patterns, might have to make even larger cuts in service than we observed under the demonstration. These substantially greater cuts might lead to adverse outcomes.

Agencies in a National System Might Not Behave Like Demonstration Agencies

The voluntary nature of the demonstration raises the possibility that participating agencies may not have been representative of home health agencies nationwide. However, we found that impacts did not differ with agency characteristics (predemonstration practice pattern, profit status, size, auspice, and status relative to cost limits) that might be expected to influence agency behavior under prospective payment. Because agency characteristics did not affect our findings, we have somewhat more confidence that our results can be generalized to all agencies.

Demonstration agencies were aware that they were being scrutinized as subjects in a demonstration. They may have responded by increasing their efforts to maintain the quality of care, and prospectively paid agencies might have made special efforts to do so. If agencies behaved in this manner, we may have underestimated the impacts of prospective payment on the quality of care provided. However, agency staff informed us during our site visits that, other than learning the new demonstration quality assurance process to which control agencies also were subject, they did not change their quality assurance procedures during the demonstration.

POLICY IMPLICATIONS

Satisfaction Measures Could Be Added to Quality Monitoring System

The small increases in patient dissatisfaction resulting from prospective payment are less worrisome than any negative effects on health and function would have been. Nevertheless, they might warrant attention. Although the impacts we observed in this demonstration were quite small (differences of two to four percentage points), differences in incentives and possible differences in agency behavior under a nationwide system could have greater impacts on measures of satisfaction. Therefore, HCFA may wish to consider monitoring patient satisfaction under a prospective payment system for home health care.

Continued Monitoring of Quality Will Be Important as a National Prospective Payment System Is Implemented

This study demonstrates that prospective payment caused no detectable adverse effects on patients' health and function. This finding should reduce concern about implementing a Medicare home health prospective payment system, as mandated by Congress. However, the findings of this demonstration cannot guarantee that health outcomes will remain unchanged as agencies make additional reductions in services. HCFA has devoted extensive attention and resources to developing a home health quality assurance system, and it will be important to monitor quality as a national system is implemented.

I. THE PER-EPISODE HOME HEALTH DEMONSTRATION AND EVALUATION

The Health Care Financing Administration's (HCFA's) Per-Episode Home Health Prospective Payment Demonstration tests the extent to which prospective payment for Medicare home health services increases efficiency in service provision. Per-episode payment encourages efficiency by giving agencies the financial incentive to reduce their costs. Specifically, under the demonstration payment system, any savings generated from lower costs per episode of patient care can result in agency profit.[1] These incentives differ greatly from those under the current system of cost-based reimbursement, which does not reward cost-containment efforts.

A. QUALITY OF HOME HEALTH CARE UNDER THE DEMONSTRATION

In this report, we present findings from an analysis of demonstration impacts on the quality of home health care during agencies' first two years of participation in the demonstration.[2] Trenholm (2000) already has shown that prospective payment indeed sharply reduced service use. Compared with control agencies, treatment agencies provided an average of 24 percent fewer visits to patients in the year after admission. Most of the difference was attributable to reductions in the number of visits by skilled nurses and home health aides, although the number of therapy visits and medical

[1]Strictly speaking, only for-profit agencies earn profits. However, for brevity, we use the term "profits" in this report to refer to both surpluses generated by nonprofit agencies and profits earned by for-profit agencies.

[2]This report originally was to describe demonstration impacts on the quality of care for the entire three-year period of the demonstration, but Congress mandated in the 1999 Balanced Budget Agreement that, beginning in 1999, the Medicare program must pay for home health services under a prospective payment system. HCFA therefore requires input as soon as possible and cannot wait for collection of all the data, which would not occur until after the demonstration has been completed.

social worker visits decreased as well. Moreover, treatment agencies discharged their patients about one month sooner than did control agencies (average episode duration, 98 days versus 131 days).

The consequences of these marked declines in the amount and intensity of home health with respect to the quality of care is an important policy issue. In theory, if the mix and level of services currently provided under cost reimbursement surpass the "point of diminishing returns," reductions in the level of home health care services could maintain (or even improve) the quality of care. However, the practical policy concern is that service reductions could have adverse effects on patient care ranging from mild decreases in patients' satisfaction to substantial decrements in their health and functional status. Policymakers therefore must determine whether savings from prospective payment entail a trade-off in the quality of care, and if so, the magnitude of that trade-off.

Evidence from previous studies on the consequences of reducing home health care services is limited and mixed. In a study of post-acute home health care provided during the two weeks after hospital discharge, patients who received an amount of home health care less than recommended by expert-derived guidelines were at increased risk for adverse outcomes (Phillips 1990). Moreover, in a study that compared outcomes of home health care patients in Medicare fee-for-service with those of home health care patients in Medicare risk plans, patients in risk plans received fewer home health visits, were discharged from home health care sooner, and had worse functional outcomes after discharge (Shaughnessy et al. 1995). Conversely, the wide regional variation in number of visits and episode duration without corresponding variation in patient outcomes suggests that agencies may be able to use considerable discretion when planning the amount of home health care for a given patient (ProPAC 1996; Schore 1995; and Welch et al. 1996). Our preliminary quality impact report on the demonstration did not find evidence that prospective payment adversely affected patient outcomes (Chen and Noveck 1998). In addition, treatment agency staff at the start

of the demonstration supported the notion that some visits could be eliminated without adversely affecting patient care (Phillips and Thompson 1997).

Several factors suggest that the quality of care is unlikely to suffer significantly from the incentive for financial gain. First, in the highly competitive home health market, an agency that develops a reputation among referring providers and patients for poor quality could suffer losses in market share, revenue, and, therefore, economies of scale. Second, patient care staff, most of whom are personally and professionally committed to serving patients by providing high-quality care, are likely to resist any efforts by agency management to reduce services below a minimum level. Third, Medicare certification surveys and the independent demonstration quality assurance (QA) process are used as ongoing monitors of quality.

The remainder of this chapter gives an overview of the Medicare home health benefit and the Medicare-certified home health industry, describes the Per-Episode Home Health Prospective Payment Demonstration, and presents a brief overview of the rest of the report. A reader who is familiar with Medicare home health may wish to proceed directly to Section C of this chapter, which describes the demonstration procedures.

B. THE MEDICARE HOME HEALTH BENEFIT

Congress established the Medicare home health care benefit in 1965, when the original Medicare program was created. Home health benefits were included to offer beneficiaries with acute conditions a less intensive and less expensive alternative to inpatient hospital care. At different times since the inception of the Medicare program, the home health benefit has been modified, partly to increase access to care.

The current Medicare home health benefit covers home health services under Parts A and B; there are no deductibles or coinsurance payments.[3] To be eligible for home health benefits, a beneficiary must (1) have Medicare coverage; (2) be homebound; (3) be under the care of a physician; and (4) need skilled nursing care, physical therapy, or speech therapy on a part-time or intermittent basis.[4]

Coverage under the home health benefit broadened considerably after the settlement of a lawsuit against HCFA in 1989, which, in turn, contributed to dramatic growth in Medicare home health expenditures after that time. Medicare spending for home health care rose from $2.5 billion in 1989 to $16.8 billion in 1995, more than tripling the share of total Medicare outlays spent on home health (U.S. General Accounting Office 1998). Nearly all this growth has been the result of an increase in the provision of services, which has coincided with a dramatic expansion in industry size. In 1989, there were roughly 5,700 Medicare-certified home care agencies; by 1997, this figure had risen to more than 10,000.

As the use of home health services has grown nationwide, service use patterns have continued to differ strikingly across regions. For example, agencies provided an average of 47 (approved)

---

[3]In some rare instances, such as when an individual does not qualify as homebound or does not have Part A coverage, home health may be covered under Medicare Part B. In cases of ambulatory Part B home health, deductibles and copayments do apply. In addition, under the Balanced Budget Act of 1997, the long-term use of home health services was transferred to Part B. Our analysis predates this change, however.

[4]Skilled nursing services are covered as long as they (1) have been ordered by a physician, (2) are required on a part-time or intermittent basis, (3) require the skills of a registered nurse (or of a licensed practical nurse or licensed vocational nurse under a registered nurse's supervision), and (4) are reasonable and necessary to treat an illness or injury. Physical therapy and speech therapy are covered if a physician's assessment recommends these services. Beneficiaries who need only occupational therapy are entitled to benefits only if they have established a prior need for skilled nursing care, speech therapy, or physical therapy in the current or preceding certification period (see Teplitsky and Janson 1985-1992, p. VII.23, Section 204.4).

visits during an episode of care to beneficiaries admitted to home health in 1990 and 1991, with the mean episode lasting 94 days. However, the average number of visits per episode varied from a low of 28 in the Pacific region to a high of 95 in the East South Central region, and the respective mean episode durations ranged from 60 days to 180 days (Schore 1995). Likewise, in 1994, Medicare home health recipients received an average of 66 visits per year nationally, but the average varied regionally from 45 per year in the Pacific region to 106 per year in the East South Central region (Health Care Financing Administration 1996). By 1997, agencies were providing an average of 73 visits per year nationwide, but ranging from a low of 32 visits per year in Washington State to a high of 161 visits per year in Louisiana (Health Care Financing Administration 1999).

The dramatic growth in Medicare home health expenditures, combined with substantial regional variation in service use and recent investigations of industry fraud and abuse, prompted Congress to legislate changes to the Medicare home health benefit. The Balanced Budget Act of 1997 originally mandated the implementation of per-episode prospective payment for Medicare home health by 1999, but the Omnibus Consolidated and Emergency Supplemental Appropriations Act of 1999 changed the implementation date to October 1, 2000. Other changes to the home health benefit under the Balanced Budget Act include:

- Reducing the per-visit cost limit from 112 percent of the mean cost to 105 percent of the median cost for freestanding agencies in the region. The Section 223 limits were frozen for reporting periods that began between July 1, 1994, and June 30, 1996.
- Defining the maximum payment for an agency by using a new algorithm that is based on annual per-beneficiary costs or per-visit costs in a base year, whichever is lower. This algorithm is commonly referred to as the Interim Payment System (IPS), because it is intended to last only until the mandated prospective payment system takes effect.
- Eliminating coverage for blood drawing when it is the only home health service required
- Redefining "part-time" and "intermittent" care

- Redefining service location on the basis of the location of the patient rather than of the agency
- Requiring additional billing information (specifically, an identifier for admitting physician and visit duration)

Demonstration agencies were exempt from the changes in the per-visit cost limit and the IPS payment algorithm during the demonstration. However, the agencies were subject to the other changes.

C. THE PER-EPISODE DEMONSTRATION

Because a cost-reimbursed payment system does not contain any mechanism that enables home health agencies to earn profits by cutting costs, it also does not offer agencies an incentive to provide services efficiently. Moreover, by reimbursing costs up to allowable limits, the system effectively subsidizes inefficient providers. The intent of per-episode prospective payment is to increase efficiency, by using the opportunity to generate profits as the primary incentive.

Ninety-one Medicare-certified home health agencies in five states--California, Florida, Illinois, Massachusetts, and Texas--enrolled in the three-year per-episode demonstration.[5] Of the 91 agencies, 47 were randomly assigned to the treatment group and received per-episode payment. The remaining 44 were assigned to the control group and continued under cost reimbursement. One control agency subsequently transferred into the treatment group near the start of the demonstration,

[5]Reflecting the United States more generally, the use of Medicare home health varied considerably in the five states. In 1997, the average annual number of visits provided per beneficiary was as follows: California, 47; Illinois, 47; Florida, 70; Massachusetts, 89; and Texas, 134 (U.S. General Accounting Office 1998). In a future report, we will compare the demonstration agencies with those nationwide on a wide variety of characteristics.

leading to a revised total of 48 treatment agencies and 43 control agencies.[6] Each agency entered the demonstration and began implementing prospective payment at the start of its fiscal year. The first entrants to the demonstration began implementing prospective payment in June 1995; the latest entrants began in January 1996. Demonstration operations continued through December 1998. (The demonstration was extended for treatment group agencies until October 1, 2000, when Medicare home health prospective payment will be implemented nationally.)

Mathematica Policy Research, Inc. (MPR) is the contractor responsible for the process and impact evaluation of the demonstration. Several other organizations also participated in the demonstration. Abt Associates, Inc. (Abt), the implementation contractor, was responsible for recruiting demonstration agencies, monitoring the status of demonstration operations, and calculating certain statistics required to determine agency payments. Palmetto Government Benefits Administrator (PGBA) was the fiscal intermediary and had responsibility for reviewing claims of and payment to treatment and control agencies. The Center for Health Policy Research (CHPR) at the University of Colorado designed and implemented a QA system for the demonstration agencies.

1. Demonstration Payment and Incentives

HCFA developed the Home Health Prospective Payment Demonstration to assess whether the profit motive can increase efficiency in the provision of Medicare home health care and thereby reduce public expenditures, without sacrificing access to care or the quality of care. Phase I of the demonstration, which tested per-visit prospective rate setting, gave agencies an opportunity to

[6]The switch was made at the request of the agency. It already had established a network with two other agencies that were assigned to the treatment group. The three agencies plan to merge fully after the demonstration has ended.

generate profits (and avoid losses) by reducing per-visit costs.[7] The evaluation of Phase I found that some agencies earned profits but did little to contain overall costs because they had no incentive to limit the services they provided. The current phase of the demonstration, Phase II, tested a per-episode prospective payment system. This system encouraged agencies to reduce their per-visit costs and provided the additional incentive to reduce visits per patient episode. Thus, the opportunity for cost savings was much greater under the Phase II payment system than under the Phase I system.

a. Payment

Agencies selected for the treatment group received a lump-sum payment for the first 120 days of home health care, regardless of the number or cost of visits provided.[8] The agencies were thus "at risk" for the costs of care incurred during this period. Only after the 120-day at-risk period and a subsequent 45-day gap in services had elapsed could an agency receive a new per-episode payment for a given Medicare beneficiary. For each visit after the 120-day at-risk period that did not begin a new episode (referred to as the "outlier period"), treatment agencies received a fixed payment rate that varied by the type of visit. In the demonstration, a treatment agency was also paid on a per-visit basis for visits made to patients admitted before the agency began demonstration operations ("phase-

[7]Although the per-visit demonstration was implemented in the same five states as is this one, most of the agencies participating in the per-episode demonstration did not participate in the per-visit demonstration. (Only agencies in the per-visit control group were eligible.) For details on the per-visit demonstration, see Brown et al. (1995).

[8]Durable medical equipment, nonroutine medical supplies, and Part B ambulatory home health services continued to be reimbursed at cost throughout the demonstration. In addition, if an agency had not provided one or more of the six home health care services covered by Medicare during the base year but began to do so during the demonstration, then those visits were also reimbursed at cost during the demonstration, as were the costs of care for which Medicare is a secondary payer. The six Medicare home health services are (1) skilled nursing, (2) physical therapy, (3) occupational therapy, (4) speech therapy, (5) home health aide, and (6) medical social work.

in" visits), and for those made to patients admitted within 120 days of the end of demonstration operations in that agency ("phase-out" visits).

Agencies selected for the control group received payments based on the cost-based system in place at the start of the demonstration. Specifically, control agencies were reimbursed for their actual per-visit costs, up to 112 percent of the mean cost incurred by all agencies (for the agency's mix of visits) in the same geographic area.

b. Prospective Rate Setting

Prospective (per-episode) rates for the at-risk period were based on a treatment agency's costs and episode profile in the fiscal year preceding its entry into the demonstration (the base year), adjusted in each demonstration year for both inflation and changes in case mix.[9] The episode profile is the average number of visits the agency provided during an episode, calculated for each of the six types of visits covered by Medicare. Payments for outlier, phase-in, and phase-out visits were also based on the agency's base-year per-visit costs (adjusted for inflation).[10] HCFA's market basket was used to adjust both the per-visit and per-episode rates for inflation.

The case-mix adjuster classified each patient into 1 of 18 groups on the basis of 12 variables that described the patient's characteristics. From this information, an aggregate case-mix index was created for each agency. At the end of each year of the demonstration, an agency's case-mix index for that year was compared with its case-mix index in the base quarter (the last quarter of the base

[9]For more information on payment rates under the demonstration, see Phillips et al. (1995).

[10]Complete data for neither episode profiles nor settled cost reports were available for a given year until some months after that year had ended. Therefore, the demonstration used preliminary initial lump-sum and per-visit rates, which were revised as final base-year data became available.

year). If the agency's case mix differed from its base-quarter case mix, its aggregate payment was retrospectively adjusted (see Trenholm 2000 for additional details).

c. Profit Sharing and Loss Sharing

To counteract the incentive to dramatically reduce services at the expense of quality and to prevent agencies from realizing excessive surpluses at public expense, HCFA shared in any profits above selected levels. If the total of a treatment agency's per-episode and per-visit prospective payments was greater than the costs for services covered by these payments, then any resulting profit greater than five percent of total allowable costs for these services was subject to profit sharing. HCFA's share was 25 percent for profits between 5 and 15 percent of total allowable costs, and increased by an amount that varied with the demonstration year for profit rates greater than 15 percent.[11]

HCFA provided a loss-sharing arrangement as a means of encouraging agencies to participate in the demonstration. It reimbursed treatment agencies for 99 percent of losses occurring during the first demonstration year, and for 98 and 97 percent of losses in the second and third demonstration years, respectively, as long as total payments were within the demonstration cost limits.

2. Other Demonstration Procedures

Other important demonstration procedures involved the process of medical review, the methods for billing, and the requirements for QA. All three procedures strongly affected treatment agencies, whereas control agencies were affected principally by the QA requirements. In this section, we briefly discuss each procedure.

---

[11]Total allowable costs were costs up to 112 percent of the mean costs incurred by all agencies (for the agency's mix of visits) in the same geographic area at the beginning of the demonstration.

a. Medical Review

PGBA conducted a limited medical review (known as an "abbreviated" medical review) of the care that treatment agencies delivered during the at-risk period of the patient episode. This review, which was limited to the admission bill, sought to determine whether the patient met the coverage criteria for home health care, and whether the agency provided at least one visit that met these criteria. As a condition of payment, PGBA required the agency to submit either HCFA 485 and 486 forms (which contain information on the patient's health and eligibility status, as well as on the home health plan of treatment) or clinical notes for admissions that coincided with an episode eligible for prospective payment. The abbreviated medical review was based on these materials. All services paid for under per-visit rate setting were subject to the usual focused medical review, under which a sample of claims was reviewed to ensure that each visit was medically reasonable and necessary.

Although all episodes initially were subject to abbreviated medical review, Medicare reduced the proportion to 25 percent in May 1996. It made this change because the claims process took longer than expected, and because 100 percent medical review was considered unrealistic under a national program.

Medical review for control agencies continued under the current (nondemonstration) regulations. For some control agencies, however, PGBA's review procedures may have differed from their predemonstration fiscal intermediaries' procedures. In these cases, the control agencies had to adopt some minor procedural changes.

b. Billing

Treatment agencies had to submit an admission bill to PGBA in order to initiate an episode of care. If a treatment admission claim was accepted (subsequent to abbreviated medical review), then

the per-episode payment was provided as a lump sum.[12] Treatment agencies also had to submit interim bills for services provided during the remainder of the at-risk period, although they received payments only for supplies.[13] For services provided near the start of the outlier period, the agencies had to "split" their standard bill into two bills--one for services provided up to the end of the at-risk period, and one for services provided from the start of the outlier period.

An agency was required to submit a discharge bill to terminate the episode of a patient who was discharged during the at-risk period or during the outlier period. PGBA would not initiate a new episode for that patient unless the previous episode had been terminated. In addition, before initiating a new episode, PGBA determined that the 120-day risk period and a 45-day gap had elapsed.

Periodic interim payments (PIPs), a means of "smoothing" cash flow for home health agencies, were originally discontinued for treatment agencies. However, to address cash flow problems in some agencies, PGBA later introduced a similar periodic payment system (called "biweekly interim payments").

Control agencies continued to submit bills under cost reimbursement and continued to be eligible to receive PIPs. PGBA based the amount of the PIP on the agency's average cost for each type of visit. For some agencies, this method differed from the practices of their predemonstration fiscal intermediaries, which based the PIP on overall agency average cost per visit. The new method

---

[12]If the admission claim was denied, interim claims for that episode were suspended for 65 days, to give the treatment agency time to file an appeal. If an appeal was filed, interim claims were suspended until a decision on the appeal for the admission claim was made. If an admission claim was denied but an appeal was not filed within 65 days, or if the denial of the admission claim was upheld on appeal, then suspended interim claims were released for possible payment under the agency's per-visit rates.

[13]In addition, agencies had to submit interim bills to receive payment during outlier periods, and to correctly calculate costs for profit or loss sharing with HCFA.

may have led to minor differences in the PIP amounts that some control agencies received before and after entering the demonstration.

c. Quality Assurance

All demonstration agencies were required to collect and submit patient-specific information to CHPR, the demonstration QA contractor. The QA procedures followed a continuous quality improvement approach. Visiting staff from demonstration agencies were required to collect information (primarily on functional status and medical condition) at admission and at discharge, or 120 days after admission, whichever came first. Similar information was also collected before admission to an inpatient facility for stays of 48 hours or longer, and on a patient's return to home health care after such an inpatient stay. CHPR used this information to develop profiles describing patient outcomes at each agency. The profiles were given to the demonstration agencies to help them improve the quality of care they provided.

II. DATA

Demonstration impacts on the quality of care could potentially manifest themselves in areas ranging from satisfaction to mortality. To capture the full spectrum of outcomes, we therefore relied on four different data sources: (1) Medicare claims data, constructed from Medicare claims and home health agency billing data; (2) quality assurance (QA) data; (3) four-month survey data, from a telephone survey of patients conducted four months after their admission to home health care; and (4) eight month survey data, from a survey of the same patients conducted eight months after admission to home health care. The claims data measure mortality and admission to services covered by Medicare Part A (hospital, skilled nursing facility (SNF), and home health care); the QA data measure home health agency staffs' assessments of patients' functional status and health services use; and the survey data measure patients' self-reports of health, functional status, and satisfaction.

Each of the four data sources offers different strengths and weaknesses for the evaluation of demonstration impacts on the quality of care. The claims data have the advantages of large size and, because outcomes are ascertained independently of treatment status, unbiased treatment-control comparisons. However, the outcome measures in the claims data--mortality and use of Medicare Part A covered health services--may be insensitive to the quality of home health care; that is, home health care quality may not affect these outcomes. The outcomes of the QA data, which consist of staff assessments of patient function and medical symptoms, may be more responsive measures to changes in the quality of home health care. Unfortunately, treatment-control comparisons in the QA data are difficult to interpret because ascertainment of outcomes may depend on treatment and control status. As described more fully in Section B, the follow-up assessments of patients that the

QA data describe were made at discharge, and prospective payment did lead treatment agencies to discharge their patients earlier than did control agencies. Thus, relative to control agency patients, treatment agency patients tended to receive their follow-up assessment at an earlier point in their course of illness. Although we cannot determine whether this differential timing of follow-up assessment introduces any bias into comparisons of treatment and control groups, bias is certainly possible. Finally, the strengths of the survey data are their measurement of patients' satisfaction with care, self-perceived health, and self-reported function, and their ascertainment of outcomes at a fixed point independent of treatment-control status. The drawback of the survey data are their small size, which limits their power.

Each data source was unique in its construction and characteristics. Sections A through D of this chapter describe in detail the analysis files created from the data sources, which are given a summary overview in Table II.1.

A. MEDICARE CLAIMS DATA

We constructed the Medicare claims data file from home health agency billing data and Medicare Part A claims data. We began by extracting claims data from the Medicare Standard Analytic File in May 1998. Because claims generally are included in the Standard Analytic File within four months after the date of service, our file allows us to track demonstration patients' use of Medicare Part A services from the time of admission to home health care through December 31, 1997.

We used data from Uniform Billing (UB-92) bill record files obtained from Palmetto Government Benefits Administration (PGBA), the demonstration fiscal intermediary, to identify

TABLE II.1

ANALYSIS SAMPLES FOR THE EVALUATION OF IMPACTS OF PROSPECTIVE PAYMENT ON THE QUALITY OF CARE

File Characteristics	Claims Data	QA Data	Four-Month Survey Data	Eight-Month Survey Data
Number of Observations				
Treatment	50,380	28,742	1,053	958
Control	45,563	20,045	1,109	925
Time Period of Data[a]	July 1995 - August 1997	May 1996 - August 1997	January 1997 - August 1997	January 1997 - August 1997
Data Sources Used in File Construction	Claims, UB-92, ARF, Abt, (QA)[b]	QA, claims, UB-92, ARF, Abt	Four-month survey, claims, UB-92, ARF, Abt	Eight-month survey, four-month survey, claims, UB-92, ARF, Abt
Main Types of Outcome Measures	Mortality; same-body-system admissions to hospital, SNF, or home health[c]	Agency staff assessments of functional status, symptoms, emergency health care use	Patient self-reports of satisfaction with care, function, and health	Patient self-reports of function and health

Claims = data from Medicare's Standard Analytic File and enrollment database; QA = quality assurance data collected by the demonstration quality contractor, Center for Health Policy Research; UB-92 = home health agency billing data from the demonstration fiscal intermediary, Palmetto Government Benefits Administrator; ARF = Area Resource File; and Abt = data on demonstration agencies collected by the demonstration implementation contractor, Abt Associates, Inc.

[a] The time period is the period during which patients were admitted to home health care. We have a complete record of Medicare Part A health services use in the claims data through December 31, 1997. The surveys were administered roughly four and eight months after admission to home health.

[b] The QA data are in parentheses because we used QA control variables only for the subgroup analyses of the claims data. The availability of the QA control variables for the claims data was poor because many of the episodes in the two data sources did not overlap chronologically.

[c] A "same-body-system admission" to each type of provider was an admission for which the admitting diagnosis involved the same general body system as did the original problem necessitating home health care (see Section A.1 of this chapter).

home health episodes as defined by demonstration rules.[1] We scanned the UB-92 files, beginning with each agency's enrollment in the demonstration, to identify the first admission for an individual and the complete set of that person's subsequent bill records. The first demonstration admission determined the episode start date from which we constructed the first record for the patient. To determine the end of the initial episode and the start of any subsequent episodes, we tracked bill records until we observed a 45-day gap in care that began after the end of the at-risk period (that is, after the first 120 days from admission). This procedure was followed regardless of whether the agency discharged and readmitted a patient within the first 165 (120 + 45) days after the initial admission. If we observed a readmission for a patient after 165 days, and a 45-day gap in care had taken place, we created a second record for the patient corresponding to this second demonstration episode. We then repeated the process until we had constructed, for each patient, a series of records for all the episodes beginning in the agency between its demonstration start and August 31, 1997.

1. Claims Outcome Variables

We merged the episodes identified from UB-92 bill record data to the Standard Analytic File, which provided information on Medicare Part A health services use, and to Medicare enrollment database files, which provided information on dates of death. The outcome variables were mortality and "same-body-system admission" to a hospital, SNF, or home health care (see Table II.2). A same-body-system admission was an admission for a diagnosis involving the same-body-system as the condition for which the patient originally was admitted to the demonstration home health agency. The coding manual of the *International Classification of Diseases, 9th Revision* (ICD-9), groups

[1]We used the UB-92 bill records rather than the claims data submitted to the Health Care Financing Administration because the UB-92 records contain data on patient characteristics at admission, which we then used to construct several key control variables.

TABLE II.2

OUTCOME VARIABLES FROM THE CLAIMS DATA

Admission for a "Same-Body-System Diagnosis" to:[a]
Hospital
Skilled nursing facility
Home health agency[b]
Mortality

NOTE: Each outcome is measured at three time points after home health admission: (1) 120 days, (2) eight months, and (3) one year.

[a]Admission for a "same-body-system diagnosis" was an admission for a diagnosis involving the same body system as was the condition for which the patient was originally admitted to the demonstration home health agency. The coding manual of the *International Classification of Diseases, 9th Revision* (ICD-9) groups ICD-9 diagnosis codes into 13 body systems (for example, "Diseases of the Circulatory System"). An admission to hospital, skilled nursing facility, or home health care was counted as an outcome only if the principal or first additional ICD-9 diagnosis for that admission fell into the same body system as the principal or first additional ICD-9 diagnosis for the original home health admission.

[b]Treatment agencies could not initiate another episode during the first 165 days after an admission. Thus, during the initial 165-day period for both treatment and control episodes, we counted only same-body-system admissions to home health agencies other than the original admitting agency. After the 165-day period, however, we counted any same-body-system home health admission, including ones to the original admitting agency.

diagnosis codes into 13 body systems (for example, "diseases of the circulatory system"). An admission to the hospital, SNF, or home health care was counted as an outcome only if the principal or first additional ICD-9 diagnosis for that admission fell into the same body system as the principal or first additional ICD-9 diagnosis for the original home health admission. We examined the occurrence of outcomes over three periods: (1) 120 days after admission (the at-risk period), (2) eight months after admission, and (3) one year after admission.

Creating the outcome variables for admission to home health care required some additional specification. Treatment agencies could not initiate another episode during the first 165 days after an admission. Therefore, for the initial 165-day period for both treatment episodes and control episodes, we counted only same-body-system admissions to home health agencies other than the original admitting agency. However, after the 165-day period had ended, we counted any same-body-system home health admission, including admissions to the original admitting agency.

2. Claims Sample

From the initial file of 119,358 records, we excluded a number of records to arrive at the final analysis file. First, we dropped 2,293 records from the three agencies that had left the demonstration soon after its start. These agencies either had been purchased by another agency or had merged with one, and the new ownership was not interested in participating in the demonstration. We also dropped an agency that had no admissions, leaving 87 agencies (46 in the treatment group and 41 in the control group). We then dropped 1,183 episodes for which we observed no home health visits. Another 1,695 episodes were dropped because data, primarily UB-92 remarks (patient characteristic) variables, were missing. These data had been inadvertently erased because of an error in the software used by PGBA. We also dropped 5,350 records of patients who had belonged to Medicare health maintenance organizations (HMOs) on or after the episode start date. We took this step

because the services provided to managed care patients were not subject to either prospective payment or traditional cost reimbursement. We dropped another 2,910 patients for whom Medicare was a secondary payer on or after the episode start date. As in the case of managed care patients, the services these patients received were not subject to per episode payment, so that the behavior of the providing agencies may have been largely independent of the demonstration incentives.

To avoid a possible bias in estimating demonstration impacts, we then dropped records of patients who already had had a demonstration episode within the previous year (9,973 records). We dropped these records because the previous discharge by prospectively paid agencies might have led to increased likelihood of readmission to home health care, which, in turn, might have affected demonstration impact estimates. Consider, for example, two identical patients, one admitted to a control agency and the other to a treatment agency. The control agency does not discharge its patient until day 300 after admission. However, the treatment agency discharges its patient on day 100. Possibly owing to the treatment agency's relatively earlier discharge, the patient is readmitted to the treatment agency for days 260 to 300. Consequently, this patient contributes two observations to the treatment group. Moreover, the second observation may be of the patient at a later stage of illness, so the likelihood of an outcome (for example, functional improvement, satisfaction, mortality, or hospitalization) may be quite different than for the original home health admission. To assess the sensitivity of the results to excluding previously admitted patients, we reanalyzed key outcomes after restoring the patients to the sample.

Finally, we restricted the file to patients on whom we had one full year of follow-up data as of the file cut-off date of December 31, 1997. We were thus left with a final analysis file of 65,284 observations.

B. QUALITY ASSURANCE DATA

The demonstration QA contractor designed and implemented a patient-outcome–based quality monitoring and continuous improvement system for the demonstration. All demonstration agencies were required to collect and submit QA information to the Center for Health Policy Research (CHPR). The QA data collection instruments are modified versions of CHPR's full Outcome Assessment System Information Set (OASIS; Shaughnessy et al. 1995). CHPR calculated agency-level, risk-adjusted profiles of patient outcomes from the QA data, which were then regularly "fed back" to the agencies to help them improve the quality of care (Shaughnessy et al. 1995). Most demonstration agencies had implemented QA data collection by May 1996, although a few did not do so until later.

A record in our QA file consists of at least two assessments of a patient's health made by agency nurses over the course of the episode. The nurses made the first assessment at the initial admission to the home health care agency and recorded the information in a QA start-of-care instrument. The nurses made the final assessment at whichever of the following occurred first: (1) discharge, or (2) 120 days after admission, and recorded the information in a follow-up/discharge instrument. For patients who required transfer to an inpatient facility for 48 hours or more, the nurses also completed interim assessment forms. An interim form was filled out based on the last home health visit just prior to the transfer, and another interim form was filled out if the patient resumed home health care from the agency following discharge from the inpatient facility.

1. Quality Assurance Outcome Variables

The QA outcome variables fall into two broad categories: (1) health measures, and (2) emergency services use. Table II.3 provides a summary list of the outcome variables.

TABLE II.3

OUTCOME VARIABLES FROM THE QUALITY ASSURANCE DATA

Health Measures
Improvement or Stabilization in:
Pain
Pressure ulcer count
Most problematic pressure ulcer
Surgical wound status
Dyspnea
Urinary tract infection
Urinary incontinence or catheter present
Confusion
Frequency of behavior problems
Grooming
Bathing
Toileting
Transferring
Ambulating
Light meal preparation
Housekeeping
Management of oral medications
Emergency Services Use
Reported Emergency Visit to:
Hospital emergency room
Outpatient clinic or "urgent care center"
Physician's office

a. Health Measures

We studied 17 health measures in the QA data, including basic activities of daily living (ADLs) and instrumental activities of daily living (IADLs) (for example, bathing, grooming, dressing, eating, transferring, management of oral medications, and light meal preparation) and symptoms (for example, pain interfering with activity, dyspnea, and confusion). The nurses scored the items at each assessment, on an ordinal severity scale. For example, the item "How often does pain interfere with the patient's activity/movement" has a four-level response: (1) none of the time, (2) some of the time, (3) most of the time, and (4) all of the time. Binary change variables of "improvement" or "stabilization" in each item were calculated from the initial admission and final follow-up scores. Improvement in a measure has the value of one if a patient's score on the scale for that item improved on the second assessment, and zero otherwise. By definition, patients who started the home health episode at the highest level of a measure were excluded from the sample for improvement in that measure, because they would not have been able to improve further. Stabilization in a measure has the value of one if a patient's score on the scale for that measure did not worsen on followup (that is, the score remained the same or improved). Patients who started the episode at the worst level of a measure were excluded from the sample for stabilization in that measure, as their score could not worsen. Thus, the number of patients "eligible" for improvement or stabilization varied for each measure.

b. Reported Emergency Health Services Use

On the basis of reports by the patient, family, or other providers, agency nurses recorded on the QA follow-up/discharge instrument any emergency visits to hospital emergency rooms, physicians' offices, outpatient clinics, or freestanding urgent care centers since the time of the last completed instrument. In addition, we took care to include emergency visits recorded on any interim

instruments (that is, instruments that were completed prior to or following a hospital stay but that were not the final instruments).[2]

2. Quality Assurance Sample

After removal of the records from the three agencies that dropped out of the demonstration, the original file from CHPR contained 102,852 unique episodes, created from pairs of start-of-care and follow-up/discharge instruments. One treatment agency and one control agency were not represented because they failed to submit adequate QA data to CHPR, leaving 86 agencies (40 in the control group and 46 in the treatment group).

We then applied additional exclusion criteria. We excluded 5,168 records of patients for whom the first assessment was made at a resumption of home health after hospitalization or for whom the final assessment was made at the last visit immediately preceding a hospitalization. These patients had been assessed at different points in their courses of illness than had those observed at initial admission to home health care, at routine discharge from home health care, and at the end of the at-risk period.[3] After dropping an additional 27,160 episodes with start dates occurring after August 31, 1997 (the cut-off date for home health admissions in our Medicare claims file), we merged the file with the claims file and dropped another 13,940 records that did not match to a claims record.[4] We then excluded 486 episodes that lasted one day or less. Finally, as for the claims data (see

---

[2]This is because the question for both interim and final assessment instruments asked the nurse if the patient had used emergency care services since the last time the instrument was completed.

[3]These observations may be useful for quality improvement, the original purpose of the QA data, but we decided to exclude them for the evaluation of the demonstration.

[4]We required an exact match to the patient's Health Insurance Claims (HIC) number, the agency's provider number, and the start-of-care date. Of the records in the merged file, 9,095 did not match the HIC number, and the remaining 4,845 did not match the start-of-care date.

Section A.2), we dropped 7,311 patients who had had previous demonstration home health admissions in the year preceding their admission. We took this step to avoid the potential bias of including more readmitted treatment agency patients who were at a different stage of illness than first-time admissions. The result was an analysis file of 48,787 episodes.[5]

C. FOUR-MONTH PATIENT SURVEY DATA

We conducted a patient telephone survey as part of the demonstration evaluation. A sample of patients admitted to demonstration home health agencies was surveyed by telephone roughly four months after home health admission, and data were collected on health status and satisfaction with care. The four-month interview coincided approximately with the end of the 120-day at-risk period covered by per-episode payment, the time at which we would most likely have observed adverse effects of prospective payment on the quality of care resulting from the incentive to reduce visits during this period. We sampled, by agency, from all admissions to demonstration agencies during their second demonstration year and the early part of their third demonstration year. We chose this period to minimize demonstration start-up or phase-out effects on agency behavior. Patients were sampled with targets of a minimum of 15 and a maximum of 100 completed *eight*-month surveys per agency. If a patient could not be interviewed, a proxy respondent was interviewed in his or her stead. Appendix A describes the fielding of the survey in detail.

---

[5]After the analysis was completed and a draft report prepared, we identified a small fraction of cases that were ineligible for demonstration payment but had not been excluded due to oversight. These were primarily cases in which Medicare was the secondary payer or the patient was a member of a particular type of managed care organization for which claims appear on Medicare files. Excluding these cases and re-estimating all of the analyses produced essentially identical results as the ones, presented in Chapter IV, in which these cases had not been excluded.

1. Four-Month Survey Outcome Variables

We analyzed survey responses in two areas: (1) satisfaction with agency care, and (2) self-reported health and functional status. Rather than ask proxy respondents for patients who had died or were in comas about the patients' health or functional status, we asked about satisfaction with agency care. Responses were dichotomized into binary variables. For example, we compared the likelihood of treatment agency patients versus control agency patients reporting they were extremely dissatisfied or mostly dissatisfied with agency care, or reporting they were able to get out of bed or out of a chair independently. We treated "don't know" or "refused" answers as missing and excluded those cases from analysis. There were no meaningful differences in item nonresponse rates between treatment and control group patients.

Several of the satisfaction items offered a range of response (for example, "strongly agree," "agree," "disagree," and "strongly disagree"). For these questions, we considered the effects of alternative cut-off points to avoid overlooking treatment-control differences in the intermediate responses. Where results were similar regardless of the cut-off point used in the definition of the outcome, we present only one result. Where a different definition yielded a different result, we present both results.

For items on ADLs and IADLs, we asked patients who reported that someone usually assisted them with an activity whether they *could* have performed the activity in the absence of help. If we had measured only whether a patient actually performed an activity, the comparison between treatment and control agencies might have been distorted. The distortion could have arisen because control agencies were more likely than treatment agencies to be providing services at four months, and patients receiving assistance might have had fewer opportunities to perform activities

independently, thus appearing relatively more impaired. The outcome variables of the four-month survey are summarized in Table II.4.

2. Four-Month Survey Sample

Of the 88 agencies remaining after exclusion of the 3 agencies that had dropped out, 2 agencies had too few admissions during the sampling period to reach the minimum targeted number of admissions, leaving 86 agencies. Two of the 86 agencies were too small to analyze individually. We combined their 12 patients into a single treatment agency because the agencies had the same owner, were located close to each other, and were both treatment group members. Thus, the final sample contained 85 agencies (44 in the treatment group and 41 in the control group) and included episodes with start dates occurring between mid-January 1997 and late August 1997, and survey dates occurring from early May 1997 to early January 1998.

We completed 2,699 four-month interviews. Of these, we dropped the following 174 cases: (1) 67 that had incorrect HIC numbers or claims denied by PGBA; (2) 29 that had started home health care after the cut-off point for start dates in our Medicare claims data (that is, after August 31, 1997); (3) 65 that had a greater than 14-day discrepancy between the survey data and the Medicare claims data in the date of admission to home health care; and (4) 13 that had no home health visits in their Medicare claims, were missing the remarks fields from their UB-92 data, or were from agencies that had dropped out of the demonstration. We were thus left with 2,525 cases (1,217 in the control group and 1,308 in the treatment group). As with claims and QA data, we dropped

TABLE II.4

OUTCOME VARIABLES FROM THE FOUR-MONTH SURVEY

Overall Satisfaction
- Satisfied with Care Received from Agency
- Of Those Satisfied, Percentage Extremely Satisfied with Care
- Would Recommend Agency to Friend or Family
- Satisfied with Current Personal Care Arrangements

Discharge
- Discharged Too Soon
- Needed Home Services After Home Health Discharge (Services Not Available, a Major Problem)

Staff
- Arrived More than Three Hours Late or Failed to Arrive
- Rushed Through Work
- Encouraged Independence
- Provided Reassurance and Emotional Support
- Paid Attention to Patient

Aide
- Completed All Work
- Did Not Come Often Enough

Nurses and Therapists
- Did Not Come Often Enough
- Were Careful and Thorough in Examination and Treatment
- Provided Sufficiently Long Visits
- Gave Clear Explanations of Medical Conditions and Treatment
- Provided Excellent Education About Care

General Health
- Reported Health Good or Excellent
- Any Days in Bed Within Past Two Weeks
- Satisfied with Life

Feeding
- Fed Self Independently
- Could Have Fed Self Independently[a]

Transferring
- Transferred from Bed or Chair Independently
- Could Have Transferred Independently

Ambulating
- Ambulated Independently
- Could Have Ambulated Independently

Bathing
- Bathed Independently
- Could Have Bathed Independently

Medications
- Took Medications Independently
- Could Have Taken Medications Independently

[a]Patients were asked both whether they usually *performed* an activity and whether they *could* have performed the activity, had there been no one to help. If we had measured only whether a patient actually performed an activity, the comparison between treatment and control agencies might have been distorted because control agencies were more likely to still be providing services at four months.

episodes for patients who had had previous demonstration home health admissions in the year preceding the index admission (453 patients), to reach a final analysis file of 2,072 observations.[6]

The results we present for analyses on the four-month data include all 85 agencies. Observations in the main analysis were weighted to represent agencies equally. We therefore were initially concerned that observations from agencies with few observations would carry very large weights, adversely affecting statistical precision through the design effect. We considered dropping agencies with fewer than nine observations for any given outcome (the sample sizes and number of observations per agency varied for each outcome, so we would have had to drop different agencies for each outcome). We did not take this step, however, because the results did not differ appreciably regardless of whether we dropped or retained agencies with fewer than nine observations.

D. EIGHT-MONTH SURVEY

The patients surveyed in the four-month survey were re-interviewed roughly eight months after their home health admission. The eight-month survey analysis file was created by appending the eight-month responses to the four-month survey analysis file.

The eight-month survey repeated the items from the four-month survey on health and functional status. However, because the majority of patients had been discharged from home health care by the time of the second survey, that survey did not include items on satisfaction with the care provided by the home health agency. From the 2,072 patients in the four-month sample, a total of 1,883 eight-month interviews were completed (958 in the treatment group and 925 in the control group). The largest value of the weights necessary to give each agency equal representation in the

[6]As in the case of the QA sample, there was a small percentage of cases in the survey sample that should have been excluded because Medicare was the secondary payer or the patient was a member of a managed care organization. We have also redone all of the analyses after dropping these cases from the survey sample and find no differences in the estimates.

analysis was 3.6, allaying our concern that observations from small agencies would receive excessively large weights, thereby distorting the results. Thus, we also included all 85 agencies for the eight-month analysis. As with the four-month survey, proxy respondents for patients who had died or were in comas were not asked additional questions about the patients' health and functional status.[7]

E. CONTROL AND SUBGROUP VARIABLES

The relatively small number of agencies randomized and differential attrition of control and treatment agencies raised the possibility that the baseline characteristics of the treatment and control groups might have differed. We thus included in our models a standard set of control variables to adjust for any such differences. The control variables, listed in Table II.5, measure characteristics at the patient, agency, and area level.

1. Standard Control Variables for Patient Characteristics

It was important to control for patient characteristics because individuals who were more severely ill, were more functionally impaired, or had certain diagnoses could have had a higher likelihood of suffering adverse outcomes, such as dissatisfaction with care, admission to hospital, functional decline, or death. We drew the standard set of control variables on patient characteristics from three data sources: (1) the "remarks" fields from the home health UB-92 bills, (2) the Medicare enrollment database, and (3) the Medicare Standard Analytic File. In the remarks field for the first UB-92 bill submitted after a demonstration admission, all agencies were required to submit information on patient characteristics needed to determine the 18-category Home Health Utilization

[7]Again, re-estimating the eight month regressions using the smaller sample in which the patients ineligible for lump-sum payments had been dropped from the four month sample made no material difference in the results.

TABLE II.5

STANDARD CONTROL VARIABLES FOR THE MAIN MULTIVARIATE ANALYSIS, BY SOURCE

Patient Level		Agency Level		Area Level
Patient Characteristics at Admission[a]	Medicare Service Use Before Demonstration Admission[b]	Base-Quarter Patient Service Use	Agency Characteristics[d]	Area Characteristics[e]
Age	Length of Preadmission Inpatient Stay (if Any)	Agency Practice-Pattern Index[c]	Chain Member	Physicians per 10,000 (1994)
Gender	Whether in Skilled Nursing Facility Within 14 Days Before Admission		Hospital Based	Nursing Home Beds per 100 Elderly Residents (1991)
Race			Proprietary	Hospital Occupancy Rate (1993)
Original Reason for Entitlement	Number of Home Health Visits in Six Months Before Admission		Small Agency (<30,000 Visits in Base Year)	
Whether Has Cancer			State	
Whether Has Diabetes	Total Part A Medicare Reimbursement in Six Months Before Admission		Rural	
Whether Has Decubiti				
Whether Needs Complex Wound Care				
Whether Has Limitations in Activities of Daily Living				
Whether Admitted to Home Health from Hospital				
Whether Has Medicaid				
Date of Admission				
Agency's Year in Demonstration[d]				

[a]From Medicare enrollment database; UB-92 remarks.

[b]From Medicare Standard Analytic Files.

[c]The agency predemonstration practice-pattern index is the case-mix–adjusted ratio of the average number of visits the agency made to its patients in the 120 days after admission during its predemonstration base quarter to the average number made by other demonstration agencies. The index is calculated from data in the Abt Associates, Inc. (Abt) base-quarter case mix file (see footnote 5 in this chapter).

[d]From base-year cost reports and Abt enrollment file.

[e]From Area Resource File.

Group case-mix adjuster. The characteristics included measures of impairment in ADL and whether the patient had certain medical conditions (cancer, diabetes, stroke, or decubitus ulcers) and certain care needs (for complex wounds).

We obtained basic patient demographic information, including the patient's age (at the start of the home health episode), gender, race, and original reason for Medicare qualification (for example, age or disability) from the Medicare enrollment files. From the Standard Analytic File, we constructed measures of Medicare service use to reflect the patients' severity of illness, including measures of both recent acute illness (whether admitted from hospital and length of previous hospital stay) and longer-term illness (home health and hospital use in the six months preceding home health admission). In the case of beneficiaries younger than 65.5 years at home health admission, we used the mean value for beneficiaries between 65.5 and 66 years of age as a proxy measure, because those younger than age 65.5 would not have been eligible for Medicare service for the complete six-month period. Medicare claims data also were incomplete for patients who had been enrolled in a managed care plan or for whom Medicare was a secondary payer for the six months preceding home health care. We therefore used the overall mean for these patients. We also included in the model three binary variables (one for each of the three types of missing data) indicating that we had imputed a missing value by substituting a mean.

Because agency behavior may have changed over time in response to changes in federal legislation and in the health care industry, we used dummy variables to indicate the period (the quarter of each calendar year) during which patients were admitted to home health care. Finally, we included an indicator variable for whether the patients' admitting home health agency was in its first versus its second or third year in the demonstration. By chance, more treatment group agencies

entered the demonstration late than entered early, and agencies could have benefited from a "learning curve" as they gained experience with the incentives of prospective payment.[7]

2. Standard Control Variables for Agency Characteristics

The characteristics of home health agencies had the potential to influence the mix of home health services delivered or the types of patients served. For example, relative to freestanding agencies, hospital-based agencies might have served more acutely ill patients. Therefore, to control for any treatment-control differences in agency characteristics, we used data from base-year Medicare cost reports and data collected by the demonstration implementation contractor, Abt Associates, Inc. (Abt). The Medicare cost reports provided information on the agencies' base-year characteristics, including for-profit status, affiliation (hospital based or freestanding), and size (small: fewer than 30,000 total visits, or large: more than 30,000 visits). The data collected by Abt during demonstration recruitment provided information on the agencies' chain membership, rural location, and practice patterns.

An agency's practice pattern could have affected patient outcomes either directly, as a measure of home health care intensity, or indirectly, as a reflection of agency and area characteristics as they existed before the demonstration start (for example, agency ownership and local market characteristics). The agency predemonstration practice pattern variable is the case-mix–adjusted ratio of the average number of visits the agency made to its patients during the 120 days after

[7]Neither the time variables nor the time-by-treatment interaction variable was significant in any of the models, suggesting a lack of any "learning curve effect" or secular time trends on treatment-control differences.

admission during its predemonstration base quarter to the average number made by other demonstration agencies.[8] A practice-pattern index value greater than one indicates that, controlling for differences in case mix, during the quarter preceding the demonstration, an agency provided more visits during the 120-day period than did all other demonstration agencies.

3. Standard Control Variables for Area Characteristics

We also controlled for area-level characteristics that might have influenced the outcomes under study. For example, the likelihood of admission to home health care or SNF might have depended on the local supply of nursing home beds. We used the 1996 Area Resource File of the Health Resources and Services Administration (which reports data from previous years) to determine the number of physicians per 10,000 residents, the number of nursing home beds per 100 elderly residents, and hospital occupancy rates.

---

[8]The practice pattern index is calculated as follows. Let subscript *i* refer to the service type and let subscript *j* refer to the case-mix cell. Using the case-mix adjuster developed by Abt, we classify an agency's patients into 1 of 18 case-mix cells ($j = 1,..,18$). Abt collected case-mix data for each agency's predemonstration base quarter. Within each case-mix cell, we multiply the average visits of each type for a given agency (n_{ij}) by its national cost limit in the base year (w_i) and then sum across the visit types. This sum essentially reflects a weighted count of the average visits for an agency within a case-mix cell. We then use these weighted counts to construct the ratio of the agency's average number of visits received by patients in the case-mix cell ($\Sigma_i w_i n_{ij}$) to the average number of visits by all agencies for this cell ($\Sigma_i w_i N_{ij}$). Finally, we arrive at a practice-pattern for each agency by summing across the 18 case-mix ratios, weighting each ratio by each agency's proportion of episodes in the case-mix cell during the base quarter (p_j). Thus, for each agency, the index practice pattern is given by:

$$\sum_j p_j \left(\frac{\sum_i w_i n_{ij}}{\sum_i w_i N_{ij}} \right).$$

4. Patient-Level Control Variables, by Data Source

The patient-level control variables included in the analyses of each of the four data sources varied somewhat by the data source. Variations arose either because we augmented the standard control variables with control variables available only within that data source, or because we dropped certain control variables for statistical reasons.

a. Claims Data

The claims outcomes focused primarily on the use of health services. Therefore, we included the following two additional control variables from the claims data: (1) the patient's total number of hospitalizations during the six months preceding his or her home health admission, and (2) the patient's total number of SNF admissions during that period.

b. Quality Assurance Data

To conduct analyses of the QA data, we included additional baseline variables in the QA data that would predict the QA outcomes of improvement or stabilization in function, symptom severity, and emergency care use. These variables, from the admission assessments by home health agency staff, measured patient behavior, mood, availability of caregiver help, and prognosis (Table II.6).[9]

We had problems with "overfitting" or "separation" of small samples or uncommon outcomes. Overfitting and separation are numerical problems that occur when data are spread over too many cells, and that are manifested by unrealistically large standard errors for the parameter estimates of some control variables (Hosmer and Lemeshow 1989). Therefore, to limit the number of regressors, we dropped some of the standard patient-level control variables for uncommon outcomes and for

[9]We did not use the additional QA control variables for the main analyses of the claims and survey data because their availability for these two data sources was poor. First, many of the episodes in the QA and claims data did not overlap chronologically. Second, agencies complied in varying degrees with the QA contractor's request to complete and submit the QA instruments.

TABLE II.6

ADDITIONAL CONTROL VARIABLES FROM THE QUALITY ASSURANCE DATA

Symptoms and Conditions at Admission
High-risk factors[a]
Medically unstable
Expresses feelings of depression
Observed to be depressed
Demonstrates cognitive impairment
Prognosis at Admission
Likelihood that treatment can be taken over[b]
Prognosis is good/fair
Life expectancy less than six months
Rehabilitative prognosis is good
Availability of Informal Care at Admission
Live-in informal help
Paid help or residing in assisted-living residence

[a]Has any of the following risk factors: heavy smoking, obesity, alcoholism, or drug dependency.

[b]By the patient, or by relatives, friends, neighbors, or paid helpers of the patient.

outcomes based on small samples. In the case of stabilization of surgical wounds, for example, the sample consisted of 9,218 patients. (Only patients with surgical wounds at baseline that were not classified as at the worst level were eligible for this outcome.) Of the 9,218, only 213, or 2.3 percent of the total, did *not* stabilize.

We dropped the following variables, which had a less than 5 percent incidence in the sample: admission by a nonurban agency (5 percent), having Medicare as a secondary payer during the six months preceding home health admission (2.8 percent), membership in an HMO during the six months preceding home health admission (2.4 percent), and coverage by Medicare for less than six months (1.5 percent). In addition, we did not include the following UB-92 remarks variables as control variables in the QA models, because they represented QA outcomes, which already had been calculated as changes from baseline: stage 3 and stage 4 decubitus ulcers, and impairment in bathing, toileting, and transferring. Finally, we dropped the number of SNF admissions during the six months preceding home health admission, as it did not correlate with the QA outcomes. No dropped variables were found to be statistically different between the treatment and control groups in the claims-based analysis, except for SNF admissions, and SNF admission rates were very similar for the two groups.

c. Four-Month and Eight-Month Surveys

For the analysis of the survey data, we included from the survey itself additional control variables important for the survey outcomes of satisfaction and function. These variables were whether the patient (1) was the respondent for the interview, (2) received a non-Medicare homemaker/aide visit during the month preceding home health admission, (3) was married at the time of home health admission, (4) received unpaid help from friends or family during the month

preceding admission, and (5) had friends or family living in the patient's household during that month.

Because of the relatively small sample sizes and our concern over overfitting and separation, we used parsimonious regression models for the survey analyses. We combined the two variables of having Medicare as a secondary payer during the six months preceding home health admission, and membership in an HMO during the during the six months preceding home health admission, into a single control variable. We did not include coverage by Medicare for less than six months as a control variable because of its low incidence. Among possible control variables measuring prior health service use, we used the single variable of total Part A Medicare reimbursement in the previous six months, rather than also including the number of hospitalizations or SNF admissions in the six months before home health admission.

5. Subgroups

As discussed more fully in Section A.1 of Chapter V, we defined a number of agency- and patient-level subgroups. The agency-level subgroups were formed by using agency-level control variables listed in Table II.5 and described in Section E.2 of this chapter. The five subgroups were (1) high or low predemonstration practice pattern (agencies with an agency predemonstration practice pattern variable above or below the median), (2) for profit or nonprofit, (3) small or large size (fewer than or more than 30,000 total visits), (4) hospital based or freestanding, and (5) above or below base-year cost limits.

The three patient-level subgroups were (1) whether the patient was able to take oral medications independently, (2) whether the patient had access to caregiver help, and (3) whether the patient had characteristics associated with a need for costly care. The first two subgroups were created from the corresponding baseline QA variables. The third subgroup was derived from a combination of several

additional QA variables.[10] As we have noted, however, we could not obtain QA variables for the complete claims sample, so the patient-level subgroup analyses were performed only in the subset of the claims records that matched to the QA data (see footnote 6 earlier).

F. SUMMARY STATISTICS FOR CONTROL VARIABLES

We present the means of patient-level control variables in Table II.7, and of agency- and area-level control variables in Table II.8. This presentation enables the reader to compare the values of the variables in the four data sources and in the treatment and control groups, and to obtain an overview of patients and agencies in the demonstration. In the tables, observations are weighted to give each agency equal representation.[11] As discussed in Section E.4, we did not use all control variables in the analysis of every data source.

1. Overview of Data Sources

The samples are comparable across all measured characteristics. Changes in case mix across time or chance variation in random sampling could have led to systematic differences across the

---

[10]Patients' expected costliness of care was based on QA variables for planned home health treatments and on UB-92 remarks variables for medical conditions/needs. The QA variables were (1) administration of insulin, (2) administration of other injections, (3) management and evaluation of patient care plan, (4) dysphagia treatment, (5) bed bath, and (6) personal care. The UB-92 remarks variables were (1) diabetes, (2) severe pressure ulcers, and (3) need for complex wound care. These variables were identified by Phillips et al. (1992) as increasing costs for the average patient by more than 30 percent. Trenholm (2000) provides a detailed discussion of the definition of the patient-expected-cost subgroup.

[11]The significance levels for the tests of equality between treatment and control group means in Tables II.7 and II.8 do not account for design effects due to the clustering of patients within agencies. Because we are describing only our analysis samples, without generalizing to the universe of all agencies, there is no need to account for clustering effects. However, we do account for the design effects associated with our use of sample weights. In our impact analyses, in contrast, we do account for design effects from both the use of weights and the clustering of patients within agencies. Sections B and C of Chapter III present a discussion of the effects of clustering and weighting in our impact analysis.

TABLE II.7

WEIGHTED TREATMENT AND CONTROL GROUP MEANS FOR PATIENT-LEVEL INDEPENDENT VARIABLES USED IN THE ANALYSIS OF CLAIMS, QUALITY ASSURANCE, FOUR-MONTH SURVEY, AND EIGHT-MONTH SURVEY SAMPLES

Variable (Source of Variable--See Key at End of Table)	Claims		QA		Four-Month Survey		Eight-Month Survey	
	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group
	(All Values are Percentages Unless Otherwise Noted)							
Demographic Measures								
Original Reason for Medicare: Reached Age 65 Years (EDB)	83.3	81.2***	83.6	81.7***	83.5	82.1	83.3	82.5
Age (Years; EDB)								
Younger than 65	8.5	9.3**	8.2	9.5**	8.7	9.5	8.8	8.9
65 to 74	28.1	29.7***	28.0	29.6***	28.4	27.6	29.0	27.8
75 to 84	40.0	39.6	40.8	40.0***	39.5	39.7	39.2	39.6
85 or older	23.5	21.4	23.0	20.9***	23.4	23.2	23.0	23.7
White (EDB)	80.6	80.7	82.0	82.1	81.8	77.9**	81.8	77.9**
Female (EDB)	63.3	63.9	64.2	63.5	65.4	63.7	66.6	64.9
Has Medicaid Buy-In for Part A and B Medicare (EDB)	26.4	23.7***	25.3	21.7***	27.5	22.8***	27.8	23.4**
Enrolled in Medicare for Less than Six Months Before Home Health Admission (EDB)	1.5	1.5	--	--	--	--	--	--
Enrolled in HMO at Some Time in Six Months Before Home Health Admission (EDB)	0.9	1.1	--	--	--	--	--	--
Had Medicare as Secondary Payer at Some Time in Six Months Before Home Health Admission (EDB)	0.3	0.3	--	--	--	--	--	--
Either in HMO or Medicare Secondary Payer at Some Time in Six Months Before Home Health Admission (EDB)	--	--	--	--	4.5	3.9	4.5	4.1
Patient Was Survey Respondent (Survey)	--	--	--	--	53.4	53.5	53.7	51.2
Medical Conditions, Symptoms, and Needs at Home Health Admission								
Cancer (UB92)	12.8	13.0	12.5	12.1	11.5	11.6	9.7	8.5
Diabetes (UB92)	20.5	20.9	19.5	19.7	19.7	20.0	19.9	19.5
Cerebrovascular Accident (UB92)	14.7	14.5	13.5	13.8	15.3	12.9*	15.9	12.8**
Stage 3 or 4 Decubitus Ulcer	4.2	4.0	--	--	3.2	3.5	2.8	3.4

TABLE II.7 *(continued)*

Variable (Source of Variable--See Key at End of Table)	Claims		QA		Four-Month Survey		Eight-Month Survey	
	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group
Need for Complicated Wound Care (UB92)	7.1	6.9	7.5	7.2	6.2	7.4	--	--
Functional Limitations (UB92)								
Bathing	72.1	71.3	--	--	71.5	69.4	70.6	68.4
Eating	27.1	27.4	23.6	23.3	26.3	24.5	25.4	22.7
Dressing	60.9	63.0***	59.0	60.5***	61.1	61.8	60.1	60.7
Toileting	36.8	37.9*	--	--	35.1	33.0	34.0	31.4
Transferring	49.8	50.9*	--	--	50.4	48.3	49.8	47.4
Has Risk Factors[a] (QA)	--	--	18.8	22.5***	--	--	--	--
Medically Unstable (QA)	--	--	49.7	54.5***	--	--	--	--
Depressed Feelings (QA)	--	--	19.3	24.3***	--	--	--	--
Displays Depressive Behaviors (QA)	--	--	20.1	24.4***	--	--	--	--
Demonstrates Disruptive Behaviors	--	--	22.0	24.7***	--	--	--	--
Prognosis at Home Health Admission (QA)								
Likelihood that Treatment Can Be Taken Over[b]	--	--	71.9	68.0***	--	--	--	--
Prognosis Is Good/Fair	--	--	87.9	84.4***	--	--	--	--
Life Expectancy Less than Six Months	--	--	7.5	6.7***	--	--	--	--
Rehabilitative Prognosis Is Good	--	--	71.0	67.1***	--	--	--	--
Availability of Informal Care at Home Health Admission								
Live-In Informal Help (QA)	--	--	44.9	46.2***	--	--	--	--
Paid Help or in Assisted-Living Residence (QA)	--	--	16.9	14.9***	--	--	--	--
Was Married at Time of Home Health Admission (Survey)	--	--	--	--	35.9	39.7	35.7	38.8
Had Non-Medicare Homemaker or Aide in Month Before (Survey) Home Health Admission	--	--	--	--	13.3	12.6	13.1	12.3

TABLE II.7 *(continued)*

Variable (Source of Variable--See Key at End of Table)	Claims		QA		Four-Month Survey		Eight-Month Survey	
	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group
Had Nonpaid Help from Friends and Family in Month Before Home Health Admission (Survey)	--	--	--	--	54.5	56.6	52.1	56.3
Had Nonpaid Friend of Family Living at Home in Month Before Home Health Admission (Survey)	--	--	--	--	33.1	36.3	31.7	34.9
Measures of Patient's Prior Service Use Before Home Health Admission								
Was in Hospital Before Home Health Admission (Claims)	36.7	39.0***	38.2	41.6***	41.3	39.9	41.4	40.1
Length of Inpatient Stay Ending in Two Weeks Before Home Health Admission (Days; Claims)	3.6	4.2***	3.7	4.2***	3.9	4.1	3.9	4.1
Whether in SNF During Two Weeks Before Home Health Admission (Claims)	17.6	16.0***	18.8	17.3***	20.6	16.1***	21.4	15.7***
Home Health Visits from Nondemonstration Agencies in Six Months Before Home Health Admission (Number; Claims)	12.7	11.9	11.8	10.4***	9.3	10.8	8.4	10.4
Hospitalizations in Six Months Before Home Health Admission (Number; Claims)	0.9	1.0***	0.9	1.0***	--	--	--	--
Whether Hospitalized in Six Months Before Home Health Admission (Claims)	--	--	--	--	66.1	66.0	66.6	66.1
SNF Admissions in Six Months Before Home Health Admission (Number Claims)	0.29	0.27**	--	--	--	--	--	--
Total Medicare Part A Reimbursement in Six Months Before Home Health Admission (in $1000s; Claims)	--	--	--	--	12.1	11.8	12.1	11.7
Time of Home Health Admission								
Quarter Admitted to Home Health Care (Claims)								
Third quarter of calendar year 1995	2.0	3.6***	--	--	--	--	--	--
Fourth quarter of calendar year 1995	4.9	6.2***	--	--	--	--	--	--
First quarter of calendar 1996	17.5	17.8	--	--	--	--	--	--
Second quarter 1996	16.2	14.1***	14.3	12.6	--	--	--	--
Third quarter 1996	13.7	13.3	19.6	19.0*	--	--	--	--
Fourth quarter 1996	13.1	12.7	19.3	18.8	--	--	--	--
First quarter 1997	13.0	13.0	18.7	18.8	24.7	30.5***	24.2	30.6
Second quarter 1997	11.9	11.7	17.6	18.4**	50.5	49.7	50.9	50.4
Third quarter 1997	7.7	7.7	10.5	12.4***	24.9	19.9***	24.9	19.0

TABLE II.7 *(continued)*

Variable (Source of Variable--See Key at End of Table)	Claims		QA		Four-Month Survey		Eight-Month Survey	
	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group
Admitted in Agency's Second or Third Demonstration Year (Abt)	38.1	40.2***	56.6	64.7***	100	100	100	100
Patients (Number)	34,518	30,766	28,742	20,045	1,053	1,019	958	925
Agencies (Number)	47	41	46	40	44	41	44	41

NOTE: Not all control variables were used in the analysis of every data source (see Section E.4 of this chapter). The dash indicates that control variables were not used in the analysis of a particular data source. Observations are weighted to represent agencies equally, and *p*-values are adjusted for the effects of weighting. Patients with a previous demonstration home health admission within the year preceding the index admission have been excluded.

QA = quality assurance; HMO = health maintenance organization; SNF = skilled nursing facility.

Key to Data Sources for Control Variables:

EDB = Medicare enrollment database
UB92 = Home health agency UB-92 billing data
QA = Quality Assurance data
Survey = MPR telephone survey
Claims = Medicare claims data
Abt = Demonstration data from Abt Associates, Inc., the demonstration implementation contractor

[a] Has any of the following risk factors: heavy smoking, obesity, alcoholism, or drug dependency.

[b] By patient, or by relatives, friends, neighbors, or paid helpers of patient.

*Means significantly different at the 0.10 level, two-tailed test.
**Means significantly different at the 0.05 level, two-tailed test.
***Means significantly different at the 0.01 level, two-tailed test.

TABLE II.8

WEIGHTED TREATMENT AND CONTROL GROUP MEANS FOR AGENCY- AND AREA-LEVEL INDEPENDENT VARIABLES USED IN THE ANALYSIS OF CLAIMS, QUALITY ASSURANCE, FOUR-MONTH SURVEY, AND EIGHT-MONTH SURVEY SAMPLES

Variable (Source of Variable--See Key at Bottom of Table)	Claims		QA		Four-Month Survey		Eight-Month Survey	
	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group	Treatment Group	Control Group
	(All Values are Percentages, Unless Otherwise Noted)							
Agency Characteristics (Abt)								
For Profit	47.8	51.2***	45.1	52.7	46.8	51.2**	46.9	51.2
Hospital Based	8.7	14.6***	9.4	16.3	9.1	14.6***	9.1	14.6
Member of a Chain	37.0	26.8***	35.0	23.9***	36.4	26.8***	36.4	26.8
Provided Fewer than 30,000 Visits in Base Year	34.8	19.5***	29.4	14.7	31.8	19.5***	31.8	19.5
Predemonstration Practice Pattern (Ratio)*	0.9	1.1***	0.9	1.1	0.9	1.1***	0.9	1.1
Area Characteristics (ARF)								
State								
California	26.1	22.0***	28.3	22.8	27.3	22.0***	27.3	22.0
Florida	8.7	9.8***	9.4	8.2	9.1	9.8	9.1	9.8
Illinois	13.0	22.0***	13.3	24.4	13.6	22.0***	13.6	22.2
Massachusetts	17.4	7.3***	18.8	8.2	18.2	7.3***	18.2	7.3
Texas	34.8	39.0***	30.2	36.4	31.8	39.0***	31.8	39.0
Urban Area	84.8	85.4	--	--	86.4	85.4	86.4	85.4
Number of Physicians per 10,000 Residents, 1994	22.0	21.5***	22.6	21.9***	22.3	21.5*	22.3	21.5
Number of Nursing Home Beds per 100 Residents Older than Age 65, 1991	5.1	5.2***	5.1	5.1	5.1	5.2	5.1	5.2
Hospital Occupancy Rate, 1993	62.0	61.0***	63.0	61.0***	63.0	61.0***	63.0	61.0
Mean Medicare Reimbursement per Beneficiary, 1991 (Dollars)	3,413	3,406	3,433	3,421	3,414	3,406	3,415	3,406

NOTE: Observations are weighted to represent agencies equally, and *p*-values are adjusted for the effects of weighting. Patients with a previous demonstration home health admission within the year preceding the index admission have been excluded.

QA = quality assurance.

Key to Data Sources for Control Variables:
Abt = Demonstration data from Abt Associates, Inc., the demonstration implementation contractor.
ARF = Area Resource File.

*An index of the case-mix–adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

*Means significantly different at the 0.10 level, two-tailed test.
**Means significantly different at the 0.05 level, two-tailed test.
***Means significantly different at the 0.01 level, two-tailed test.

sources, greatly complicating the interpretation of estimated demonstration impacts in the different data sources. However, systematic differences did not arise.

2. Patient Characteristics

Demonstration patients were quite vulnerable and frail. Roughly two-thirds of the patients were women, and two-thirds were aged 75 years or older. One-quarter were poor enough to qualify for Medicaid. Most patients suffered multiple limitations in ADLs and from chronic illnesses (diabetes in 20 percent, a history of stroke in 14 percent, and a history of cancer in 13 percent). Many patients also were recovering from acute illnesses, with roughly 40 percent entering home health care from a hospital stay, and as much as 20 percent having been in an SNF during the 14 days preceding this care. The QA data indicated a high prevalence of baseline affective and cognitive symptoms (with nearly one-quarter of patients having depressed feelings, depressive behaviors, or disruptive behaviors). Roughly half the patients were judged by agency staff to be medically unstable.

In the Medicare claims data, there were numerous differences between patients of control agencies and those of treatment agencies that were small but reached statistical significance because of the dataset's large size and statistical power. Compared with treatment agency patients, control agency patients were slightly younger but more impaired in dressing, toileting, and transferring. More control agency patients entered home health care from a hospital and fewer from an SNF (explained in part by the higher proportion of hospital-based agencies in the control group). Control agency patients had slightly longer hospital stays before entering home health care, slightly more hospitalizations during the six months preceding home health care, and slightly more admissions to SNFs. More control agency patients were admitted to home health care during the first or second quarter of the demonstration, and more were admitted during their agencies' second or third demonstration year.

The QA data, also a large dataset, likewise contained multiple small, but statistically significant treatment-control differences. Many of the treatment-control differences in the claims data were mirrored in the QA data, including age, dressing impairment (one of the ADL limitations), entry to home health care from hospital versus SNF, number of home health visits during the six months preceding the current home health care episode, hospital length of stay preceding the episode, and number of hospitalizations during the six months preceding the episode. In addition, it was more likely that control group patients than treatment group patients had risk factors, were medically unstable, had depressed feelings, displayed depressive behaviors, and demonstrated disruptive behaviors. Agency staff were also less likely to judge control group patients as having good prognoses, good rehabilitative potential, or likelihood that family or friends would take over treatment. Conversely, fewer control group patients were estimated to have a life expectancy of less than six months. More control group patients were admitted to home health care during the last two quarters for which data were collected and during their agency's second or third demonstration year.

Finally, the survey data exhibited a few significant differences. As in the claims and QA data, control agency patients, as described by survey data, were less likely to have been in an SNF before entering home health care. They were more likely to be married at admission to home health care. More control group patients were admitted to home health care during the first quarter of 1997.

3. Agency and Area Characteristics

The agency and area characteristics of the treatment and control groups differed in several important ways. The magnitudes and directions of these differences were consistent across all four data sources, although not all comparisons reached statistical significance. More control agencies than treatment agencies were for-profit entities, hospital based, and large, and more of them exhibited high-practice patterns. Fewer control agencies were members of a chain. A higher

proportion of control agencies than treatment agencies were located in Illinois, and a lower proportion were located in Massachusetts. Control agencies operated in areas with slightly fewer physicians per capita, slightly more nursing home beds per capita, and slightly lower hospital occupancy rates.

4. Implications for the Analysis of Program Impacts

The similarity of patients and agencies across the different data sources is reassuring. The findings from any single data source are thus likely to reflect the demonstration as a whole and not stem from peculiarities of that particular data set.

Although there were several treatment-control differences in the control variables, none of these differences was large enough in magnitude to raise any concern about bias. The control variables for patient characteristics indicated that control agency patients were slightly more functionally impaired, and were judged by agency staff to be "sicker." Differences in agency and area characteristics were somewhat larger. For example, the treatment agencies' mean practice-pattern index was lower than that of the control agencies. In an unadjusted comparison, these baseline differences in patient, agency, and area characteristics could lead to the erroneous conclusion that prospective payment affected the quality of home health care. The baseline differences between the treatment and control groups underscore the importance of estimating program impacts with regression models in which the differences can be controlled.[12]

[12]In a future report, we will discuss how the overall characteristics of patients, agencies, and areas affect the external validity and generalizability of our findings.

III. METHODS

To control for residual treatment and control agency differences that persist despite random assignment and to improve statistical precision, we used logit models to estimate the main and subgroup effects of prospective payment. The dependent variables in these models are measurements of patients' health status, satisfaction, or health care utilization at various follow-up points after admission to home health care. The main independent variable of interest is the treatment or control status of the patient's home health agency. The remaining independent variables are control variables that measure baseline characteristics of the patient, agency, and area at the time of admission.

A. LOGISTIC REGRESSION MODELS FOR ESTIMATING IMPACTS

All the dependent variables studied in this report are binary. We used logistic regression models to estimate demonstration effects.

1. Estimating Main Effects

The logit model for estimating program effects is:

$$(1) \quad p(Y_i=1) = \frac{e^{x_i'\beta+T\delta}}{1+e^{x_i'\beta+T\delta}},$$

where:

$p(Y_i=1)$ = the probability that the binary outcome variable Y_i for the ith observation equals one

x_i = the data vector of control variables for the ith observation, with elements $x_i' = [1\ x_{1i}\ x_{2i} \ldots\ x_{pi}]$

β = the vector of coefficients for the control variables, with elements $\beta' = [\beta_0\ \beta_1\ \beta_2 \ldots\ \beta_p]$, and β_0 is the intercept term

T = the indicator variable for treatment status and has the value of one if the observation belongs to the treatment group and zero if it is in the control group

δ = the coefficient of the treatment indicator variable T.

The coefficient δ on the variable for treatment status does not directly measure the estimated impact of the payment method. It is the log of the estimated odds-ratio of experiencing the outcome for the treatment group relative to the control group. That is, $\delta = \ln \frac{P_T/(1-P_T)}{P_C/(1-P_C)}$, where P_T is the predicted probability that an individual in the treatment group has $Y = 1$, and P_C is the analogous estimate for the same individual if he or she were in the control group. We used the p-value for the coefficient δ to test the hypothesis that the treatment effect is significantly different than zero. To estimate the demonstration impact on the probability that $Y_i = 1$, a somewhat more intuitive measure than odds-ratios, we used the coefficients estimated from the model to generate two predicted probabilities for each observation: one assuming that the observation belongs to the treatment group ($T = 1$), and the other assuming that it belongs to the control group ($T = 0$). The impact estimate is the mean difference between these estimated probabilities, across the sample.

Throughout the tables of results, we present as a point of reference the unadjusted control group mean of each outcome variable alongside the estimated impact. The unadjusted control group mean provides a reasonable estimate of the mean value for the outcome variable that might be expected to occur in the absence of the demonstration. We used this mean to assess the relative magnitude and importance of the estimated impact.

2. Estimating Subgroup Impacts

Whether certain agency types respond differently to prospective payment, and whether specific patient subgroups are affected differentially by prospective payment are important policy questions. For the subgroup analyses, we used logit models similar to equation (1). However, the exponent of

e in equation (1) now has the form $x'\beta + T\delta + \sum_{k=1}^{n}(T*s_k)\gamma_k$, where, as before, $x'\beta$ is the product of the vector of control variables with its coefficient vector and $T\delta$ is the product of the treatment status indicator with its coefficient, but $T*s_k$ are new interaction variables formed by multiplication of the treatment status variable T with the kth subgroup indicator s_k, and the γ_k are the coefficients on these interaction terms. The $T*s_k$ interaction variable has a value of one only if the observation belongs both to the kth subgroup *and* the treatment group; otherwise it has a value of zero.

We estimated impacts for the agency-level and patient-level subgroups somewhat differently. A key goal of the agency subgroup analysis was to project more accurately our impact estimates to agencies outside the demonstration. We therefore estimated impacts for agency subgroups through a single regression model that included all five treatment status-agency subgroup interaction terms. This model enabled us to pinpoint agency characteristics responsible for differential demonstration effects, net of other agency characteristics.

We were concerned that the patient-level subgroups might be so highly correlated (for example, high-cost patients and patients unable to take oral medications independently) that differential effects might be obscured in a single regression model. Furthermore, we believed the "gross" effect of selected patient characteristics was a more important policy issue than was the mechanism underlying these effects; that is, we were less interested in whether differential effects would "net out" if other patient characteristics were simultaneously controlled. We therefore estimated impacts for the three patient subgroups by using three separate regression models, each with a single treatment status-patient subgroup interaction term.

To assess whether the effect of prospective payment is uniform *across* different subgroups, we examined the statistical significance of the γ_k coefficients on the interaction terms. For example, a significant p-value for the coefficient on the interaction term of treatment status by for-profit status

indicates that the effect of prospective payment on that outcome is significantly different for for-profit and nonprofit agencies.

We also studied the effect of prospective payment *within* individual subgroups.[1] To test whether there is a statistically significant treatment impact *within* the *k*th subgroup, we tested whether $(\delta+\gamma_k)$ is significantly different than zero.[2] To use the example of the for-profit subgroup again, a statistically significant *p*-value for the test indicates that the effect of prospective payment on that outcome is significantly different for control agencies and treatment agencies *within the subgroup of for-profit agencies*.

We estimated the within-subgroup impact by setting the indicator variable for the subgroup appropriately and then obtaining the predicted value of the outcome variable for each observation, first as if it were from a treatment agency ($T = 1$), and then as if from a control agency ($T = 0$). The mean difference between these two predicted values for the sample gives an adjusted estimate of the impact of prospective payment on the probability of the outcome within each subgroup.

Because of small sample sizes, we present agency- and patient-level subgroup analyses only for outcomes obtained from claims data and QA data. Some agency-level subgroups in the survey data were extremely small, as agency characteristics were not evenly distributed across agencies.[3]

---

[1]It is not always possible to identify small or moderate impacts for subgroups, as tests of statistical significance lose power as the sample size decreases. The smaller the size of the subgroup (all else equal), the less likely we are to reject the hypothesis that the payment method has no effect, for any given true effect size.

[2]To be precise, we tested and evaluated the interaction effects at the *mean values* $\bar{s}_j$ of the *other* subgroup variables. Thus, the estimated impact of the treatment effect within subgroup s_k is $\delta+\gamma_k+\sum_{j\neq k}\bar{s}_j\gamma_j$, and the test statistic is a *t*-statistic with the estimated impact in the numerator and the variance of the estimated impact in the denominator.

[3]Take, for example, "could have eaten independently," a typical outcome in the four-month survey, for which the overall sample mean was 10 percent "no." For this outcome, there would have (continued...)

Subgroup sample sizes in the survey data for the patient-level subgroups were somewhat more balanced, but our power to detect reasonable differences was still limited, especially compared with the QA and claims data.

B. ISSUES IN WEIGHTING OBSERVATIONS

Given the complex structure of our data, with randomization and intervention at the agency level, analysis of impacts at the patient level, and patient observations clustered within agencies, sample weights are an important part of the analysis. We used one weighting strategy for the main analysis and an alternative weighting strategy as a sensitivity test of the results.

1. Weighting to Represent Agencies Equally

Our main impact analyses weight observations to give each agency equal representation, called "agency equal" weighting. We use this approach for two reasons. First, because the payment method was implemented at the agency level, the agency is the behavioral unit of interest. Second, the use of weighted data ensures that the experiences of a few large agencies do not dominate the impact estimates.

For each outcome variable to be analyzed, we constructed the agency equal weight (w_i) for the ith agency as the ratio of the average number of patient-level observations per agency (n/k) to the number in the agency (n_i):

$$(2) \quad w_i = \frac{n/k}{n_i},$$

---

[3](...continued)
been 214 observations in the hospital-based subgroup--68 in the treatment agencies (roughly 7 "no" outcomes and 61 "yes" outcomes) and 146 in the control agencies (roughly 15 "no" outcomes and 131 "yes" outcomes). Subgroup sample sizes would have been similarly small for several of the other outcomes and agency subgroups, with an even worse situation in the eight-month survey.

where n is the total number of observations across all agencies, and k is the number of agencies. Each observation from agency i was then weighted by w_i. Because n, n_i, and k all vary with the outcome, we constructed a separate set of weights for each outcome. Patient observations were thus weighted in inverse proportion to the size of their agency.

Extremely large weights (from agencies with very few observations) inflate the design effect, decrease the precision of estimated impacts, and, if a heavily weighted agency has an anomalous mean value, may distort the estimates. In the four-month survey, weights ranged from 0.21 to 5.6, in the eight-month survey from 0.24 to 3.6, and in the claims data from 0.16 to 25.9. In contrast, weights in the QA data ranged from 0.11 to 568. For 28 QA outcomes, the maximum weight exceeded 50, and of these, the maximum weight exceeded 150 in 18 outcomes. The QA weight distributions were heavily skewed to the left, with only a few extremely large weights in the tail. A typical weight distribution was that for improvement in pain, for which the 75th percentile was at 3.0, the 90th percentile at 7.8, the 95th percentile at 14, and the last 5 percent, or four weights, had values of 16.6, 20.3, 73.2, and 366.1.

We truncated the weights for the QA data (Cox and Cohen 1985). Each observation was assigned the lesser of the actual calculated weight, w_i, or a maximum allowable weight w_{max}. We determined by inspection that the 85th percentile was a natural cut-off point for all weight distributions. We then rank ordered the 16 85th percentile values for improvement outcomes (which ranged from 4.3 to 11) and the 16 values for stabilization outcomes (which ranged from 5.8 to 10.7) and selected the median value of each list of 16 as the maximum allowable weight for the entire group of outcomes (that is, 5.6 for improvement outcomes and 6.1 for stabilization outcomes). The maximum allowable weight for the separate emergency care outcomes was set at 6.0, the 85th percentile for the emergency care outcome weight distribution.

2. Robustness of Results to Alternative Weighting Schemes

Our regression model results are potentially sensitive to the sample weights we use. Although agency equal weighting might limit the influence of large agencies on the analysis, it might also accord undue influence to small agencies. In addition, the QA results might be dependent on our weight truncation strategy. To investigate the sensitivity of our results to the weighting approach, we reestimated demonstration impacts with alternative weights that represented agencies proportional to their size. Similar impact estimates under the two approaches would strongly suggest that no particular agency exerted undue influence on the analysis, and that the results are broadly interpretable for policy purposes. Conversely, although dissimilar results under the two weighting schemes do not necessarily indicate that the main ("agency-equal") results are incorrect, they do suggest that additional analysis is required to understand the reasons for the discrepancy and the implications for policy.

The claims data are the most accurate measure of agency size for calculating the "proportional to size" weights, because agencies must submit a claim for every admission in order to receive payment. For the measure of size in the claims data analysis, we simply used the number of demonstration episodes within each agency in the entire claims dataset. For measures of size in the QA and survey data, we used the claims-based counts of admissions per agency during the following periods: May 1996 to August 1997 for the QA data, and January 1997 to August 1997 for the survey data.[4] The size weight for the ith agency is then simply:

[4]The number of completed QA assessments per agency is an inaccurate measure of agency size because of agencies' variable compliance in submitting QA data to CHPR. The number of surveys per agency does not reflect agency size because we deliberately oversampled small agencies and capped the number of surveys from large agencies.

$$(3) \quad w_{i_{size}} = \frac{N_i/N}{n_i/n},$$

where N_i is the number of demonstration admissions for the agency in the claims data, N is the total number of admissions for all agencies in the claims data, n_i is the number of analysis sample observations for the agency, and n is the total number of analysis sample observations. Both N_i and N are counted over the period appropriate for the particular dataset, and both n and n_i are sample sizes for the analysis dataset.

C. STATISTICAL HYPOTHESIS TESTING

Testing the null hypothesis that prospective payment had no impacts on the outcomes requires correct estimation of the standard errors of the treatment-control coefficients. It also requires a decision about the level of significance at which to reject the null hypothesis.

1. Design Effects

To obtain the appropriate standard errors for our impact estimates, we used SUDAAN software, which is specifically designed for the analysis of clustered data. In our data, the clustering of patients within home health agencies might generate a correlation between the response variables that is due to effects of the agency on patient outcomes. Standard statistical software, which assumes simple random sampling from an infinite population, generally underestimates variances in clustered data by an amount, called the "design effect," that increases as the intracluster correlation increases.[5] Underestimation of the variances, in turn, leads to falsely small p-values, so that demonstration

[5]Our sample actually includes the entire population of patients admitted to demonstration agencies during the period of the demonstration. However, because we wished to make inferences about how outcomes would be affected for patients admitted at other times and to other agencies, we treated patients in the data as though they were drawn in a two-stage random sample from the pool of all (future) patients in all agencies.

impacts appear more significant than they really are. SUDAAN corrects for this design effect and also accounts for the increase in the standard errors caused by the use of sample weights.

2. Significance Levels

For each outcome, we used a two-tailed *t*-statistic to test the null hypothesis that there is no difference between the regression-adjusted population means for treatment and control agencies.[6] The associated *p*-value is the probability of obtaining the observed estimate under the null hypothesis. We chose a *p*-value of less than 0.10 to reject the null hypothesis and, thus, to establish that a demonstration impact is statistically significant. At this *p*-value, however, approximately 10 percent of independent tests would show, simply by chance, a statistically significant treatment-control difference in the absence of a true program effect (known as a Type I error). Therefore, in assessing whether a statistically significant treatment-control difference, especially one with a *p*-value between 0.05 and 0.10, should be interpreted as a true program impact, we also considered whether the sign and magnitude of the estimated effect were consistent with those for related outcomes.

The design effects discussed in the previous section limited our power to detect small treatment-control differences. Design effects for the data sources and outcomes in this report ranged from one to seven (including the effects of both weighting and clustering), corresponding to intracluster correlation coefficients, *k*, of 0 to 0.1. Table III.1 displays some representative minimum detectable differences for simple comparisons of treatment and control group means for the different data

[6]In general, we were concerned that prospective payment would have adverse effects on the quality of home health care and on patient outcomes. We used two-tailed tests in our analyses to avoid confusion, and to flag estimated treatment-control differences of the "wrong" expected sign that were large enough to be statistically significant. For impacts with the "correct" expected sign, a two-tailed test is less likely than a one-tailed test at the same significance level to reject the hypothesis of no demonstration effect (all else equal).

sources and shows how the minimum detectable difference increases with increasing design effect. In general, even for some of the outcomes with smaller sample sizes, we would be able to detect relatively small differences in the range of 5 to 10 percent for binary variables with a mean near .5. However, if the intraclass correlation is larger, the detectable differences for binary variables with means of .20 or smaller become quite large in proportion to the mean. Thus, impacts of policy-relevant size could be missed for relatively rare outcomes on which agencies differ markedly.

TABLE III.1

MINIMUM DETECTABLE EFFECTS FOR DIFFERENT DATA SOURCES, INCLUSIVE OF DESIGN EFFECTS

Data Source	Number of Observations per Group	Number of Agencies per Group (m)	Binary Variable Mean (Percent)	Minimum Detectable Difference (Percentage Points)[a]		
				$k = 0$	$k = 0.01$	$k = 0.1$
Claims Data						
Outcome 1[b]	32,600	44	50	1.1	3.2	9.6
			20	0.9	2.6	7.6
Subgroup (50 Percent of Agencies): Outcome 1	16,300	22	50	1.6	4.5	13.4
			20	1.2	3.6	10.7
Subgroup (25 Percent of Patients): Outcome 1	8,150	44	50	2.2	3.7	9.7
			20	1.8	3.0	7.8
QA Data						
Outcome 1[c]	23,400	43	50	1.3	3.3	9.6
			20	1.0	2.6	7.7
Outcome 2[d]	10,000	43	50	2.0	3.6	9.7
			20	1.6	2.9	7.8
Outcome 3[e]	5,000	43	50	2.8	4.1	9.9
			20	2.2	3.3	7.9
Subgroup (50 Percent of Agencies): Outcome 1	11,700	22	50	1.8	4.6	13.6
			20	1.5	3.7	10.9
Subgroup (50 Percent of Agencies): Outcome 2	5,000	22	50	2.8	5.1	13.8
			20	2.2	4.1	11.0
Subgroup (50 Percent of Agencies): Outcome 3	2,500	22	50	4.0	5.8	14.0
			20	3.2	4.7	11.2
Subgroup (25 Percent of Patients): Outcome 1	5,850	43	50	2.6	4.0	9.9
			20	2.1	3.2	7.9
Subgroup (25 Percent of Patients): Outcome 2	2,500	43	50	5.6	6.3	10.9
			20	4.5	5.1	8.7
Subgroup (25 Percent of Patients): Outcome 3	625	43	50	7.9	8.4	12.2
			20	6.3	6.8	9.7

TABLE III.1 *(continued)*

[a] Minimum detectable effects are calculated for two-tailed tests of differences in means, at the five percent significance level, with 80 percent power. For comparison of (unadjusted) means:

$$MDD = 2.8\sigma\sqrt{\frac{2k}{m} + \frac{2(1-k)}{mn}},$$

where k is the proportion of total variance in the outcome measure that is between agencies, rather than within agencies, m is the number of agencies per group (43 or 44 for the full sample, depending on the data source), n is the number of observations per agency, and σ is the standard deviation of the outcome variable, in this case $\sqrt{p(1-p)}$ for a binary variable. MDD = minimum detectable difference.

[b] The sample size for all outcomes in the claims data was equal to the full claims data sample of 65,284. There were 87 agencies in the claims data.

[c] For a QA outcome with sample size equal to the full QA data sample of 46,787. There were 86 agencies in the QA data.

[d] For a QA outcome with a sample size of 20,000.

[e] For a QA outcome with a sample size of 10,000.

IV. DEMONSTRATION IMPACTS ON PATIENT OUTCOMES

A. POTENTIAL EFFECTS ON PATIENT OUTCOMES

Research has not yet determined how substantial reductions in service provision, such as those by prospectively paid agencies, might affect patient satisfaction and the quality of care (Trenholm 2000). As discussed in Chapter I, empirical studies have reached mixed conclusions. In one study, patients who received amounts of post-acute home health care that were less than recommended in expert guidelines had an increased risk for adverse events (Phillips 1990). In another study, patients belonging to health maintenance organizations received less home health care than did fee-for-service patients and had relatively lower levels of patient functioning (Shaughnessy et al. 1994). Other studies have not found a relationship between the amount of home health care provided and mortality or use of hospital, skilled nursing facility (SNF), or additional home health care (Schore 1994; and Welch et al. 1996). In the Channeling Demonstration, patients receiving relatively more care had increased levels of satisfaction with personal care arrangements and with life in general (Kemper 1988).

Improved, unchanged, and worsened patient outcomes accompanying reductions in the amount of home health care all are consistent with theoretical considerations. Suppose, for example, that an agency responds to the incentives of prospective payment by providing "less" home health care, such as fewer home visits, earlier discharge, less frequent dressing changes, and less costly dressing supplies.[1] Furthermore, suppose the agency does not otherwise change the "content" or quality of the care. The hypothetical solid curve in Figure IV.1 (labeled "typical quality of care") relates

[1]Treatment agencies provided fewer visits and discharged patients earlier than did control agencies (Trenholm 2000). In our site visits for this study, agencies reported trying such strategies as reducing dressing changes and purchasing less expensive supplies (Phillips and Thompson 1997).

patient outcomes to the amount of home health care and shows possible effects of such a reduction. In the first instance, the amount of home care falls, from *Control* to *Prospective Payment*. Because the curve is flat, we do not observe any decrement in patient function (movement from point *a* to point *b*). In the second instance, the amount of care falls the same amount, from *Control** to *Prospective Payment**; in contrast to the first instance, however, the curve has a steep slope, so patient functioning worsens (movement from point *c* to point *d*).[2]

It is conceivable that too much home health care could actually worsen some patient outcomes (Fisher and Welch 1999). For example, too much help from agency staff could promote patient passivity and dependency. This possibility is shown by the dotted curve in Figure IV.1, where a fall in the amount of care from *Control* to *Prospective Payment* actually improves patient outcomes (movement from point *f* to point *b*).

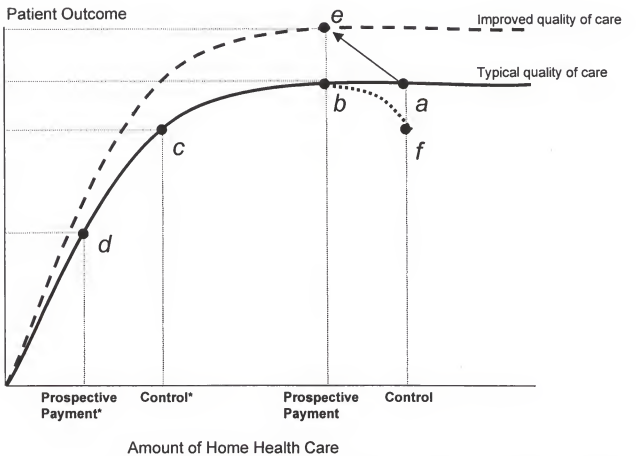
Finally, suppose agencies responded to prospective payment with innovations that simultaneously reduced the amount of care *and* improved the content and quality of care. In our site visits, treatment agencies did in fact describe strategies that potentially could accomplish both goals (Phillips and Thompson 1997):

- Improved use of staff time
  - Increased use of the telephone
  - More-critical reviews of the utility of each visit
  - Increased use of laptop computers for documentation

[2]The shapes of the curves reflect our assumptions about the relationship between the amount of home health care and patient outcomes. When the amount of home health care services is low, the curve initially slopes upward because providing additional home care services improves patients' outcomes. At larger amounts of home health care services, however, more visits produce dwindling marginal gains, so the curve flattens out.

FIGURE IV.1

CURVES OF HYPOTHESIZED RELATIONSHIP BETWEEN PATIENT OUTCOME AND AMOUNT OF HOME HEALTH CARE SERVICE



NOTE: The curves show hypothesized relationships between a patient outcome and the amount of home health care provided. If quality of care remains the same on the solid "typical quality of care" curve, a decrease in the amount of home health care provided from *Control* to *Prospective Payment* causes no change in patient outcome (movement from *a* to *b*). However, if the curve actually has a downward portion (dotted line), the decrease from *Control* to *Prospective Payment* causes the patient outcome to improve (movement from *f* to *b*). If the decrease in the amount of home health care provided goes from *Control** to *Prospective Payment**, then the patient outcome also worsens. Finally, if quality of care improves under the demonstration from the solid "typical quality of care" curve to the dashed "improved quality of care" curve, a decrease in the amount of home health care is associated with an improvement in patient outcome (movement from *a* to *e*).

- Reduction of resource use/substitution of less expensive resources
  - Intensive instruction of patients and families in self-care
  - Early involvement of community-based service providers
- Improved timing of care
  - More visits earlier in the episode
  - Improved home health discharge planning
- Increased adherence to best practices and reduction of practice variation
  - Implementation of standardized treatment protocols and care maps
  - Increased use of specialists (for example, wound care specialists)
- Increased accountability
  - Improved documentation
  - Tighter tracking of resource use and quality of care

The possibility of less albeit better care is shown by the dashed curve in Figure IV.1 (labeled "improved quality of care"). For any given amount of home care services, improved quality of care achieves better patient outcomes than does typical care.[3] With a shift to higher quality of care, a fall in the amount of home health care from *Control* to *Prospective Payment* is accompanied by improved patient outcomes (movement from point *a* to point *e*). Thus, theory does not point clearly to how the reductions in home health care service will affect patient outcomes.

[3]In our site visits, treatment agencies reported trying several new strategies, all of which had the potential to improve the quality of care: implementing standardized treatment protocols and care maps, hiring specialists (for example, wound care specialist nurses), using the telephone more frequently, using laptop computers to improve documentation, instructing patients and their families intensively in self-care, and involving community-based providers early on (Phillips and Thompson 1997).

B. IMPACTS ON HEALTH AND FUNCTIONAL STATUS

The explosive growth in home health care in recent years and the wide regional variations in the amount of home health care provided suggest that many home health care markets are already operating on the flat part of the outcome-service curve (that is, where reductions in service will not cause measurable changes in patient health and function). Indeed, we find no evidence that prospective payment adversely affected a broad range of health and functioning measures, from self-perceived health, to basic activities of daily living (ADLs) and instrumental activities of daily living (IADLs), to selected medical symptoms.

1. Self-Reported General Health and Basic Activities of Daily Living

Prospective payment had no detectable effect on any of the measures of global health. Patients in the treatment agencies and those in the control agencies reported roughly the same self-perceived health status, satisfaction with life, and likelihood of days restricted to bed at four and eight months (Table IV.1). The demonstration also had no apparent impacts on basic ADLs (Tables IV.2 and IV.3).[4] The quality assurance (QA) outcomes, which are dynamic measures of change in basic ADLs since home health admission, show that patients of prospectively paid agencies were just as likely as patients of control agencies to improve or to remain stable.

The survey outcomes, which are static "snapshots" of basic ADL abilities at four and eight months after admission, showed no consistent or plausible pattern of effects. The small but statistically significant negative estimated differences in "could eat" and "did transfer" at four months had all but disappeared by eight months (Table IV.3). Furthermore, there were no longer any

[4]See also Table IV.9.

TABLE IV.1

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN PATIENTS' SELF-REPORTS OF GENERAL HEALTH

Outcome	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (p-Value)[b]
Reported Health Good or Excellent			
At four months	1,877	43.0	-0.02 (0.99)
At eight months	1,740	46.4	-1.0 (0.7)
Any Days in Bed Within Past Two Weeks			
At four months	1,851	25.2	2.2 (0.34)
At eight months	1,713	22.1	-0.8 (0.8)
Satisfied with Life			
At four months	1,818	64.7	2.2 (0.41)
At eight months	1,682	62.9	2.9 (0.2)

SOURCE: Four-month and eight-month patient surveys.

NOTES: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

As discussed in Chapter II, we also re-estimated these survey regressions after excluding an additional 112 patients ineligible for demonstration lump-sum payment (primarily because Medicare was the secondary payer and because the patient was a member of an HMO), and the results were essentially unchanged.

[a]Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b]The p-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

TABLE IV.2

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN IMPROVEMENT AND STABILIZATION IN BASIC ACTIVITIES OF DAILY LIVING FROM QUALITY ASSURANCE DATA

Outcome	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (*p*-Value)[b]
Grooming			
Improvement	24,655	53.9	-2.2 (0.33)
Stabilization	45,218	87.3	-0.08 (0.95)
Bathing			
Improvement	37,039	51.6	-1.0 (0.67)
Stabilization	45,570	84.5	0.47 (0.68)
Toileting			
Improvement	14,807	55.1	-1.3 (0.60)
Stabilization	46,073	91.1	-0.03 (0.97)
Transferring			
Improvement	25,250	46.9	-1.8 (0.46)
Stabilization	47,660	88.1	-1.0 (0.36)
Ambulation			
Improvement	36,347	31.9	-0.9 (0.68)
Stabilization	48,653	90.2	-0.48 (0.61)

SOURCE: Quality assurance data.

NOTE: Quality assurance outcomes were assessed by home health agency nurses. The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

[a]Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b]The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

TABLE IV.3

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN BASIC ACTIVITIES OF DAILY LIVING FROM FOUR- AND EIGHT-MONTH SURVEY DATA

Outcome	When Measured	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (p-Value)[a]
Bathing				
Did	Four months	1,906	56.5	-1.5 (0.48)
	Eight months	1,742	56.1	3.4* (0.09)
Could	Four months	1,884	65.0	-0.5 (0.81)
	Eight months	1,728	65.0	3.7** (0.05)
Eating				
Did	Four months	1,880	83.5	-3.7* (0.06)
	Eight months	1,716	84.0	-4.1** (0.03)
Could	Four months	1,868	92.8	-4.1** (0.02)
	Eight months	1,710	91.7	-1.9 (0.14)
Transferring				
Did	Four months	1,912	70.3	-4.2* (0.08)
	Eight months	1,750	73.9	-1.9 (0.4)
Could	Four months	1,897	84.7	-2.0 (0.26)
	Eight months	1,737	84.8	-1.0 (0.6)
Ambulating				
Did	Four months	1,863	70.2	-3.4 (0.11)
	Eight months	1,524	85.1	-2.6 (0.23)
Could	Four months	1,849	82.4	-1.3 (0.57)
	Eight months	1,516	94.4	-0.23 (0.9)

SOURCE: Four-month and eight-month patient surveys.

NOTES: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

TABLE IV.3 *(continued)*

As discussed in Chapter II, we also re-estimated these survey regressions after excluding an additional 112 patients ineligible for demonstration lump-sum payment (primarily because Medicare was the secondary payer and because the patient was a member of an HMO), and the results were essentially unchanged.

[a] Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.

negative differences on "could" outcomes at eight months (recall that the "could" outcomes measured patients' intrinsic ability to perform activities regardless of the availability of help, whereas the "did" outcomes could be affected by the availability of help). In addition, there were small, significantly *positive* differences on "did bathe" and "could bathe" at eight months. The mixed pattern of small negative differences early in the episode versus small positive differences later in the episode, all in different ADLs, does not seem plausible. Finally, the pattern of outcomes displays no correspondence with the way in which prospectively paid agencies decreased the number of visits (that is, by reducing skilled nursing and aide visits throughout the episode and concentrating the provision of therapy visits in the at-risk period). We therefore conclude that prospective payment did not have any impacts on the basic ADL outcomes.

2. Instrumental Activities of Daily Living

Neither the QA data nor the survey data produced evidence that prospective payment had impacts on any of the measures of IADLs (Table IV.4). The treatment group and the control group did not differ significantly on any of the improvement and stabilization outcomes in meal preparation, housekeeping, or management of medications. A small relative advantage of the treatment group in "could take medications" at four months was an isolated difference that was not present at eight months, suggesting a statistical artifact.

3. Medical Symptoms and Outcomes

We find no evidence to suggest that prospective payment had a real effect on any items in this category. Most outcomes for treatment and control agencies did not differ, although we did observe small positive treatment-control differences in improvement in the level of confusion and in stabilization of urinary incontinence, and an isolated negative difference in surgical wound

TABLE IV.4

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN INSTRUMENTAL ACTIVITIES OF DAILY LIVING FROM QUALITY ASSURANCE AND SURVEY DATA

Outcome	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (*p*-Value)[b]
Quality Assurance Data			
Light Meal Preparation			
Improvement	28,682	46.6	−2.5 (0.3)
Stabilization	48,653	89.0	0.6 (0.6)
Housekeeping			
Improvement	44,083	41.5	0.21 (0.9)
Stabilization	29,585	75.7	1.8 (0.5)
Management of Medications			
Improvement	24,362	36.7	−2.3 (0.37)
Stabilization	38,168	87.2	0.7 (0.6)
Survey Data			
Did Take Medications			
Four months	1,869	52.4	1.7 (0.44)
Eight months	1,682	55.9	1.4 (0.5)
Could Take Medications			
Four months	1,846	68.9	5.0** (0.02)
Eight months	1,664	72.6	0.4 (0.9)

SOURCE: Quality assurance and patient survey data.

NOTES: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

As discussed in Chapter II, we also re-estimated these survey regressions after excluding an additional 112 patients ineligible for demonstration lump-sum payment (primarily because Medicare was the secondary payer and because the patient was a member of an HMO), and the results were essentially unchanged.

[a]Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b]The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**Significantly different from zero at the 0.05 level, two-tailed test.

improvement (Table IV.5). The number of treatment-control comparisons would lead us to expect two significant differences by chance alone at the 0.10 level. In addition, we have no reason to believe the sole negative difference represents a true impact. No literature supports this outcome as being a particularly sensitive indicator of quality of care; no difference in stabilization in surgical wounds was observed; no difference in the related outcomes of improvement in and stabilization of most problematic pressure ulcer was observed; and we know, from site visit data, that treatment agencies were actively hiring wound care specialists (specifically, enterostomal nurses).

C. IMPACTS ON USE OF HEALTH SERVICES

Previous research has failed to discern a correlation between the amount of home care and the amount of hospital and SNF use (Schore 1994; and Welch et al. 1996). Therefore, we did not expect to observe appreciable effects from prospective payment on the outcomes of emergency care use and same-body-system admissions to hospital, SNF, or home health care (Table IV.6). Instead, we find weak evidence suggesting that, if anything, prospective payment may have *reduced* the use of some of these health services.

1. Same-Body-System Diagnosis Hospitalizations and Use of Emergency Health Services

Relative to cost reimbursement, prospective payment was associated with significantly fewer hospitalizations for same-body-system diagnoses (Table IV.6). The reduction appeared at the 120-day point but did not reach statistical significance. However, it continued to increase, attaining statistical significance by eight months and by one year. By one year after home health admission, treatment agency patients were 2.4 percentage points less likely than control agency patients to have had a same-body-system diagnosis hospitalization (a reduction of seven percent relative to the control group mean).

TABLE IV.5

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN MEDICAL SYMPTOMS AND OUTCOMES

Outcome (Data Source)	N[a]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[b] (*p*-Value)[c]
Pain (QA)			
Improvement	31,462	52.2	2.2 (0.28)
Stabilization	45,496	83.1	1.1 (0.26)
Most Problematic Pressure Ulcer (QA)			
Improvement	3,844	72.6	-1.0 (0.81)
Stabilization	48,455	97.5	-0.14 (0.7)
Surgical Wound Status (QA)			
Improvement	9,774	86.2	-4.2** (0.05)
Stabilization	9,218	97.4	0.17 (0.83)
Dyspnea (QA)			
Improvement	31,908	45.3	1.4 (0.57)
Stabilization	47,307	79.2	1.8 (0.18)
Urinary Incontinence or Catheter Present (QA)			
Improvement	9,982	43.9	2.5 (0.38)
Stabilization	46,382	93.8	1.1** (0.03)
Confusion (QA)			
Improvement	15,823	37.7	5.7* (0.05)
Stabilization	43,661	84.9	2.0 (0.12)
Behavior Problem Frequency (QA)			
Improvement	5,758	62.5	-1.2 (0.72)
Stabilization	46,511	91.5	0.72 (0.42)
Mortality (Medicare Claims)			
Within 120 days	65,284	9.5	0.1 (0.85)
Within eight months	65,284	15.4	-0.2 (0.8)
Within one year	65,284	20.3	-0.7 (0.32)

TABLE IV.5 *(continued)*

SOURCE: Quality assurance and Medicare claims data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

[a] Patients who were already at the best level of a measure at the start of an episode of care could not improve in that measure. Patients who were already at the worst level of a measure at the start of care could not stabilize in that measure. Thus, the number of patients "eligible" for improvement or stabilization differed within measures and across measures.

[b] Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

QA = quality assurance data.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

TABLE IV.6

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN HEALTH SERVICES USE

Outcome (Data Source)	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (*p*-Value)[b]
Emergency Care (QA)			
Hospital emergency room	48,787	10.7	-0.9
			(0.3)
Outpatient clinic or urgent care clinic	48,787	1.3	-0.5**
			(0.05)
Physician's office	48,787	3.0	-1.0***
			(0.004)
Any of the above	48,787	13.9	-2.1*
			(0.08)
Admission to Hospital for Same-Body-System Diagnosis (Medicare Claims)			
Within 120 days	65,284	21.2	-1.1
			(0.20)
Within eight months	65,284	29.4	-2.0**
			(0.05)
Within one year	65,284	35.2	-2.4**
			(0.04)
Admission to SNF for Same-Body-System Diagnosis (Medicare Claims)			
Within 120 days	65,284	5.8	-0.5
			(0.22)
Within eight months	65,284	9.1	-0.5
			(0.38)
Within one year	65,284	11.5	-0.4
			(0.49)
Admission to HHA for Same-Body-System Diagnosis (Medicare Claims)[c]			
Within 120 days	65,284	5.6	1.1*
			(0.09)
Within eight months	65,284	11.1	1.8*
			(0.06)
Within one year	65,284	17.2	2.4**
			(0.03)

SOURCE: Quality assurance and Medicare claims data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

TABLE IV.6 *(continued)*

[a]Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b]The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[c]Because treatment agencies could not initiate another episode within the first 165 days after an admission, in the initial 165-day period for both treatment and control episodes we counted only same-body-system diagnosis admissions to agencies *other* than the original admitting agency. After the 165-day period, however, we counted any same-body-system home health admission, including ones to the original admitting agency (see Chapter II, Section C.2).

QA = quality assurance data; SNF = skilled nursing facility; HHA = home health agency.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

Prospective payment was associated with substantial reductions in emergency visits to outpatient clinics/urgent care clinics and physicians' offices (Table IV.6). There was a 0.5 percentage point drop in emergency visits to outpatient clinics/urgent care clinics (a 33 percent drop relative to the control group mean), and a 1.0 percentage point drop in emergency visits to physicians' offices (a 38 percent drop relative to the control group mean). We observe no significant impact on the use of the hospital emergency room. For the combined outcome of any emergency visit, prospective payment was associated with roughly a two percentage point reduction (15 percent relative to the control group mean).

We consider the reductions in both hospital admissions and emergency visits as a "byproduct" of a home health care payment system surprising, as even special interventions to decrease hospital and emergency room use have had mixed success (Rich et al. 1995; and Weinberger et al. 1996). The decrease in use occurred in conjunction with a cut in services by prospectively paid agencies, but without discernible changes in health status or function.

a. Possible Explanations for Observed Reductions

In this section, we consider four possible explanations for the observed reductions in hospital and emergency use. They are (1) differential treatment-control measurement of the outcome variables due to differing periods of observation, (2) improvements in the quality of care in conjunction with reductions in the number of visits, (3) a direct response to the reduction in the number of visits, and (4) unmeasured differences between treatment and control agencies.

Differential Measurement of Outcomes. This explanation says that the apparent reduction in use of emergency care services is merely an artifact of the design of the QA data collection. Patients' use of emergency care services was only recorded while they were under agency care. Due

to earlier discharge, patients of prospectively paid agencies had a shorter window of time during which emergency care events could be observed. The small observed reduction in hospital admissions, however, cannot be explained by treatment-control differences in data collection, since information on hospitalizations came from the claims data, which are recorded in an unbiased fashion.[5]

Improved Quality of Care and Fewer Visits. The second explanation accepts the reductions as real impacts of prospective payment and says that treatment agencies simultaneously improved the quality of services and reduced the quantity of services; that is, in Figure IV.1, they shifted from the "typical quality of care" curve to the "improved quality of care" curve (in this case, an improvement in patient outcomes means less hospital and emergency use). Perhaps treatment agencies' assertions, conveyed to us during our site visits--that they had improved patient and caregiver education in self-care skills, prevention of avoidable emergencies, and self-efficacy and mastery ("ego-building," in the words of one agency)--were true. However, as we discuss in Section D.2 of this chapter, patients of treatment agencies did not *perceive* their agencies as doing a better job of education, encouragement of independence, or provision of emotional support.

Additional Attention Increases Hospital and Emergency Care Use. The third explanation also accepts the reductions as true demonstration effects and says that excessive home health visits might by themselves increase emergency care and hospital use. According to this explanation, the "typical quality of care" curve in Figure IV.1 follows the dashed line downward, control agencies are on the descending limb of the curve, and treatment agencies improve patient outcomes simply by reducing services. Suppose, for example, a patient has mild clinical abnormalities that tend to

[5] Interestingly, Schore (2000), using Medicare claims data, found a small reduction in hospital emergency room use among treatment agency patients.

resolve spontaneously, so that emergency medical care provides little or no benefit. If more frequent contact by home care providers increases the probability of detecting these conditions, then the providers may make more referrals for emergency care, thereby increasing subsequent hospitalization. Fisher and Welch (1999) and Weinberger et al. (1996), in a trial of intensified posthospital follow-up and primary care, have advanced this argument for the unexpected increase in readmission rates and trend in the direction of higher mortality rates.

Two observations are consistent with this explanation: (1) the pattern of reduction in emergency services use, and (2) the pattern of reduction in hospital use. With respect to the first observation, we observe decreases in the number of visits to outpatient clinics and physicians' offices but not to hospital emergency rooms. If home health nurses in the control agencies are generating additional "optional" emergency visits, then one might expect to initially observe these visits in outpatient settings in which patients feel comfortable (for example, offices and clinics) rather than in the more intimidating or expensive hospital emergency room. With respect to the second observation (the reduction in hospital use), it occurs in hospitalizations for all causes, as would be observed as a nonspecific effect of additional attention from nurses and not from any specific improvement in quality (Schore 1999).

Unmeasured Differences. The fourth explanation says that the association between prospective payment and reductions in hospital and emergency use is a spurious one resulting from unmeasured differences between treatment and control agencies. Because the number of agencies is small, randomization may not have completely eliminated underlying treatment-control differences that are unrelated to the demonstration but that correlate with hospital and emergency use.[6] Certain agencies,

[6]The observation that all hospitalizations fell, rather than only those for same-body-system admissions, is consistent with this fourth explanation as well (Schore 2000).

for example, might tend to have a relatively high use of hospital care because of the practice patterns of referring physicians or the practice style of the immediate local community of providers. These differences might not be completely captured by all of the control variables. We were able to use claims data to investigate this possibility for the hospitalization outcomes because the base-quarter case-mix data contain information on hospital use, which applies to the period preceding the demonstration and is exogenous to it. (We could not use the QA data to assess the emergency visit outcomes because we lacked comparable data on base-period emergency visits.) Indeed, four control agencies, all of which were for-profit entities from Illinois, stood out from all other agencies as having much higher percentages of patients in the base-quarter with "embedded" hospital stays (that is, intervening hospital stays during the 120 days after home health admission).

We performed two alternative analyses to control for the possible effect of pre-existing hospitalization practices. In both, the reductions in same-body-system hospitalizations became attennated and statistically insignificant (see Appendix B). In the first analysis, we removed the four control agencies with high prior hospitalization rates. The control group mean fell from 35 percent in the full sample to 34 percent after removing the four agencies, and the impact fell from −2.4 percentage points ($p = 0.04$ in the full sample) to −1.6 percentage points ($p = 0.14$ in the smaller sample). In the second analysis, we used the full sample but explicitly controlled for the base-quarter embedded hospitalization rate in the regression. The results of this analysis were the same.

b. Cautious Acceptance of Results

The results of the main analysis lead us to cautiously conclude that prospective payment has reduced same-body-system hospitalizations and emergency visits. With the data at hand we are unable to confirm or refute these findings or to further analyze the mechanism through which they

might occur. The influence of the four control agencies with high baseline rates of hospitalization does raise the possibility that baseline agency characteristics could have contributed to the findings, and that our results are specific to this particular set of agencies. Furthermore, we do not have clear evidence to explain the process by which same-body-system hospitalizations and emergency care were reduced, so we accept the findings with some hesitation.

2. Same-Body-System Diagnosis Admissions to Home Health Agencies and Skilled Nursing Facilities

Patients of treatment agencies were more likely than patients of control agencies to have same-body-system diagnosis admissions to home health agencies, with a slightly more than two percentage point increase by one year (14 percent relative to the control group mean; Table IV.6). However, three observations suggest that the increase in admissions does not necessarily represent any decrement in quality of care due to prospective payment.

The first observation is that an atypical outlier agency in the treatment group may have skewed the treatment group mean. As detailed in Section E of this chapter, a single agency in the treatment group suffered a mishap unrelated to the demonstration that caused a large, temporary loss of a large number of the agency's patients to other agencies. Removing this agency from the analysis sample eliminated the apparent impact.

The second observation is that because treatment agency patients spent a greater amount of time during the follow-up year *out* of home health care than did control agency patients (due to their earlier discharge), they had a greater window of opportunity to be admitted by nondemonstration home health agencies. It fact, treatment and control agency patients received nearly identical numbers of nondemonstration agency visits between demonstration agency discharge and the end

of the follow-up year, suggesting that treatment agency patients were not discharged with greater home health needs (Schore 2000).

The third observation is the overall lack of adverse impacts of prospective payment on use of other health services and on health and functioning. If the greater likelihood of same body system diagnosis admission to home health care represented poor quality of care by treatment agencies, we would have expected to see some effects among those other outcomes as well.

D. IMPACTS ON PATIENT SATISFACTION

It seems intuitive that increasing the number of home health visits must eventually produce diminishing returns on improved patient health and functioning. However, one could imagine that patient satisfaction might continue to increase even with very frequent visits. The Channeling Demonstration found that higher patient satisfaction was associated with more services (Kemper 1988). Nevertheless, even patient satisfaction cannot be limitless and must therefore begin to level off at some point. On the ascending portion of the curves of Figure IV.1, then, reductions in visits will lower satisfaction, whereas on the flat parts, reductions in service will have minimal effects.

In fact, the demonstration did have negative impacts on specific areas of patient satisfaction, which we have categorized as (1) general satisfaction, (2) satisfaction with technical care, and (3) satisfaction with interpersonal care (Institute of Medicine 1990). In this section, we discuss the impacts of prospective payment on these measures.

1. General Satisfaction and Satisfaction with Technical Care

Measures of general satisfaction with agency care were high, without appreciable differences between prospectively paid and control agencies (Table IV.7). In both groups, nearly 96 percent of patients reported being satisfied with their agency, and roughly 95 percent would recommend their

agency to a friend or family who needed home health care.[7] In addition, similar proportions of patients in both groups reported extreme satisfaction with their agencies' care. Although more patients of treatment agencies had been discharged from home health care by the time of the four-month survey, comparable proportions of patients in both groups believed discharge occurred too quickly and wanted home health services after discharge that were unavailable. Both groups were equally satisfied at four and eight months with then-current arrangements for personal care. Satisfaction with technical or medical interventions (for example, nurses' and therapists' examinations and treatments, explanations of medical conditions and treatments, and education about care) were also unaffected by prospective payment (Table IV.8).

2. Interpersonal Care

In contrast, patients of prospectively paid agencies were significantly more likely to be dissatisfied with staff performance in providing interpersonal care (Table IV.8). Although roughly 90 percent of patients in both groups reported that agency staff almost never rushed through their work, the percentage of treatment group patients describing staff as rushing through work most or all of the time was approximately two percentage points higher, a nearly 50 percent increase relative to the control group mean. Similarly, the percentage of patients who strongly agreed that agency staff encouraged them to be independent and to take care of themselves was identical, but the percentage of treatment patients disagreeing or strongly disagreeing was nearly four percentage points higher, also an increase of roughly 50 percent relative to the control group mean. Finally,

[7] Patients were first asked whether they were satisfied or dissatisfied with the care received from the agency. Patients who were "satisfied" were then asked whether they were extremely, mostly, or somewhat satisfied.

TABLE IV.7

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN GENERAL SATISFACTION WITH CARE PROVIDED BY THE AGENCY

	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (*p*-Value)[b]
Overall Satisfaction			
Satisfied with Care Received from Agency[c]	2,014	96.0	0.7 (0.52)
Of Those Satisfied, Percentage Extremely Satisfied with Care Received from Agency[d]	1,804	62.3	-0.2 (0.93)
Would Recommend Agency to Friend or Family	1,999	94.9	0.7 (0.45)
Satisfied with Current Personal Care Arrangements[e]			
Four months	1,855	82.6	-1.5 (0.44)
Eight months	1,853	78.0	-1.5 (0.5)
Discharge			
Discharged Too Soon	1,439	13.9	1.5 (0.45)
Needed Home Services After Home Health Discharge--Not Available and a Big Problem	1,470	10.0	-1.4 (0.37)

SOURCE: Four-month patient survey.

NOTES: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

As discussed in Chapter II, we also re-estimated these survey regressions after excluding an additional 112 patients ineligible for demonstration lump-sum payment (primarily because Medicare was the secondary payer and because the patient was a member of an HMO), and the results were essentially unchanged.

[a] Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[c] Patients were first asked whether they were satisfied or dissatisfied with the care received from the agency. Patients who were "satisfied" were then asked whether they were extremely, mostly, or somewhat satisfied.

[d] Versus mostly or somewhat satisfied.

[e] Versus partly satisfied or dissatisfied.

TABLE IV.8

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SATISFACTION WITH AGENCY STAFF

	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (*p*-Value)[b]
Staff			
Staff Arrived Three Hours Late or More or Failed to Come at All[c]	1,967		
Little or none of the time		89.8	-1.6 (0.36)
Staff Rushed Through Work[c]	1,949		
Little or none of the time		91.1	0.3 (0.86)
Most or all of the time		4.3	2.1** (0.03)
Staff Encouraged Independence[d]	1,741		
Strongly agree		28.5	-0.04 (0.99)
Disagree or strongly disagree		7.6	3.6** (0.04)
Staff's Reassurance and Emotional Support[e]	1,923		
Excellent		40.5	-0.8 (0.74)
Staff Paid Attention to Patient[e]	1,934		
All of the time		76.8	-1.9 (0.44)
Some of the time, or little or none of the time		4.7	3.0*** (0.01)
Aide			
Aide Completed All Work	854	84.0	-3.8 (0.27)
Aide Did Not Come Often Enough	898	9.5	2.0 (0.48)
Nurses and Therapists			
Nurses and Therapists Did Not Come Often Enough	1,988	8.1	0.33 (0.85)
Nurses' and Therapists' Care and Thoroughness in Examination and Treatment[e]	1,987		
Excellent		40.7	2.3 (0.31)

TABLE IV.8 *(continued)*

	N	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[a] (*p*-Value)[b]
Nurses' and Therapists' Visits Long Enough[a]	1,919		
All of the time		76.4	2.9
			(0.24)
Nurses and Therapists Gave Clear Explanations of Medical Conditions and Treatment[a]	1,854		
All of the time		66.9	-0.6
			(0.84)
Nurses' and Therapists' Teaching About Care[c]	1,769		
Excellent		36.0	1.2
			(0.62)

SOURCE: Four-month patient survey.

NOTES: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis.

As discussed in Chapter II, we also re-estimated these survey regressions after excluding an additional 112 patients ineligible for demonstration lump-sum payment (primarily because Medicare was the secondary payer and because the patient was a member of an HMO), and the results were essentially unchanged.

[a] Estimated differences are regression adjusted through logit models to control for preexisting differences between treatment and control agencies.

[b] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[c] Response categories were: all of the time, most of the time, some of the time, and little or none of the time.

[d] Response categories were: strongly agree, agree, disagree, and strongly disagree.

[e] Response categories were: excellent, very good, good, fair, and poor.

**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

although equivalent percentages of patients reported that agency staff paid attention to what they had to say all of the time, more patients in the treatment group than in the control group believed agency staff paid attention to them only some, a little, or none of the time (the treatment-control difference of three percentage points translates into a 60 percent increase relative to the control group mean).

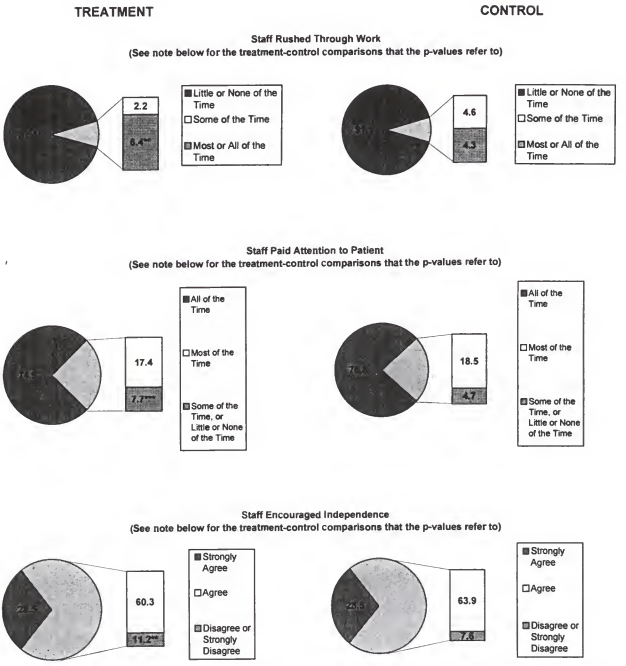
In general, the proportions of satisfied patients in the treatment and control groups were roughly the same, with the difference in the proportion of extremely dissatisfied patients explaining overall differences in satisfaction with interpersonal care. Figure IV.2 presents this effect graphically. The proportion of satisfied patients is shown in dark grey in the pie charts. The residual proportions of less than completely satisfied patients, shown in the light grey portions of the pie charts, are further expanded into the adjacent bar charts, in which the extreme category of dissatisfaction is shaded medium grey.[8] In each instance, the dark grey segments of the pie charts, representing those who were satisfied, are roughly the same in both treatment and control agencies, whereas all the medium grey portions in the bar charts, representing those who were extremely dissatisfied, are significantly larger in the treatment group. Prospective payment appears to have pushed into extreme dissatisfaction some patients who otherwise might have been only mildly dissatisfied.

---

[8]Each chart is actually based on two sets of regressions. In the first set, the outcome was the most positive response (that is, "little or none of the time" for "staff rushed through work," "all of the time" for "staff paid attention," and "strongly agree" for "staff encouraged independence"). In the second set of regressions, the outcome was the most negative response (that is, "most or all of the time" for "staff rushed through work," "some of the time/little or none of the time" for "staff paid attention," and "disagree/strongly disagree" for "staff encouraged independence"). In the control group charts, the dark grey (pie charts) and medium grey (bar charts) areas are the unadjusted means, and the light grey (pie charts) and white areas (bar charts) are the residuals to reach 100 percent. In the treatment group charts, the dark and medium grey areas are derived by adding the adjusted effects calculated from the respective regression model to the corresponding unadjusted control group means, and the light grey and white areas are the residuals to reach 100 percent. The derived areas in white are thus not tested.

FIGURE IV.2

DEMONSTRATION IMPACTS ON SELECTED OUTCOMES OF PATIENT SATISFACTION WITH INTERPERSONAL CARE, 4 MONTH SURVEY



NOTE: The possible responses for "Rushed Through Work" and "Paid Attention" were: all of the time, most of the time, some of the time, and little or none of the time. For "Encouraged Independence" they were: strongly agree, agree, disagree, and strongly disagree. For each satisfaction item, we performed two analyses: (1) defining the outcome as the most positive response(s) (little or none of the time for "rushed," all of the time for "paid attention," and strongly agree for "encouraged independence") versus all other responses; and (2) defining the outcome as the most negative response (most or all of the time for "rushed," some of the time/little or none of the time for "paid attention," and disagree/strongly disagree for "encouraged independence"). In the control group charts, the dark grey (pie charts) and medium grey (bar charts) areas are the unadjusted means, and the light grey (pie charts) and white areas (bar charts) are the residuals to reach 100 percent. In the treatment group charts, the dark and medium grey areas are derived by adding the adjusted effects from the respective regression model to the corresponding unadjusted control group means, and the light grey and white areas are the residuals to reach 100 percent. The significance tests refer only to the treatment effect in the regression model for that particular outcome. The derived areas in white are thus not tested.

**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

The patients dissatisfied with the different items tended to overlap. Of patients dissatisfied with any of the three items, roughly 60 percent were dissatisfied with one item, 35 percent with two items, and the remaining 5 percent with all three. This distribution was the same for treatment and control agency patients.

The dissatisfaction effects appeared to operate consistently across all treatment agencies. The results were robust to alternative weighting schemes (described in the next section), and examination of scatter plots does not suggest that the impacts were concentrated in a few agencies (discussed in Appendix B).

We hypothesized that staff from prospectively paid agencies might have been more rushed and more focused on technical or medical tasks to the exclusion of interpersonal behaviors, such as encouraging patients and heeding patients' opinions. In fact, other analyses of time sheet survey data indicate that visits by treatment agencies were not appreciably shorter than those by control agencies (Cheh, forthcoming). Treatment and control agencies provided visits having the same average durations, and treatment agencies were actually more likely to be in the group of agencies providing longer visits. Nevertheless, it is possible that treatment agency staff were more hurried than control agency staff, as treatment agencies provided fewer visits and discharged patients earlier. Therefore, treatment agency staff might be trying to accomplish as much in fewer visits of the same or somewhat longer duration.

We wish to emphasize that, although the *relative* impacts on dissatisfaction seem large, the *absolute* size of these effects was small. For example, because only 4.4 percent of control agency patients felt staff rushed through their work, the estimated treatment-control difference of 2.2 percentage points is proportionally large, even though nearly 93 percent of the treatment agency patients still felt staff almost never rushed through work. Similarly, the other two effects on

dissatisfaction (encouraging independence and paying attention to the patient) are absolutely small yet relatively large.

E. SENSITIVITY TESTS

We performed sensitivity tests to determine whether our estimates were contingent on our methodology or were driven by particular observations. First, we investigated the robustness of our results to weighting observations to represent agencies proportional to their size. Analyzes in which observations are weighted proportional to agency size are useful because small agencies that have anomalous results may distort the "agency equal weighted" analyzes. In the "weighted proportional to agency size" analysis, more weight is given to larger agencies. In cases in which the two weighting schemes yielded different estimates or levels of significance, we inspected agency-level means to identify both small agencies with atypical results (which could disproportionately affect the agency equal analysis) and large agencies with atypical results (which could skew the agency size analysis).

Second, we reanalyzed the data after adding back the home health episodes of patients with previous demonstration admissions, which we had excluded from the main analysis. Excluding patients with prior demonstration admissions, who are probably sicker than patients with no such admissions, might make our sample unrepresentative of home health users in general. In addition, perhaps through impacts on the quality of care, prospective payment might increase the need for readmission to home care. Analyzes that exclude patients sampled on repeat admissions might miss impacts associated with increased readmission to home health care and might fail to capture the entire effect of the demonstration. Obtaining similar results from analyzes that included patients with previous demonstration admissions and analyzes that excluded these patients would thus

strengthen our confidence that sample construction had not caused us to overlook any demonstration impacts.

1. Weighting Observations Proportional to Agency Size

The results from weighting observations proportional to agency size support the conclusion that prospective payment had negligible impacts on measures of health or functioning. In Tables IV.9 and IV.10, we summarize the analyzes of health and ADL outcomes conducted with agency weighting proportional to agency size. (The actual regression-adjusted estimated differences from these analyzes are presented in Appendix B.) Directions and significance levels of treatment-control differences are indicated with plus and minus signs for *p*-values less than or equal to 0.10, and by zeros for *p*-values greater than 0.10.[9]

Tables IV.9 and IV.10 indicate that the pattern of no sustained impacts or no consistent impacts on health and function outcomes observed in the agency equal analysis held up in the agency size analysis. The following treatment-control differences that were significant or of borderline significance in the agency equal analysis were no longer so in the agency size analysis: eating, bathing, and transferring, in the survey data (Table IV.9), and improvement in confusion, in the QA data (Table IV.10). Inspection of scatterplots of agency-level means versus agency size for these outcomes confirmed in general that upward weighting of small outlying agencies explained part of

[9]Because of the numerous statistical tests, it would have been reasonable to make an adjustment for multiple tests, such as with Bonferroni's or Tukey's correction, or to use a more stringent *p*-value, say 0.01, thereby substantially reducing the number of significant differences. However, because we were concerned about potential deleterious effects of prospective payment, we preferred making a Type I error than a Type II error. We therefore adhered to the previously chosen 0.10 level of significance.

TABLE IV.9

SUMMARY OF DIFFERENT WEIGHTING SCHEMES ON ESTIMATED EFFECTS OF PROSPECTIVE PAYMENT ON BASIC AND INSTRUMENTAL ACTIVITIES OF DAILY LIVING FROM FOUR- AND EIGHT-MONTH SURVEYS

	Equal Weight[a]		Size Weight[b]	
	Four Months	Eight Months	Four Months	Eight Months
General Health				
Health Good or Excellent	0	0	0	0
Any Bed Days in Past Two Weeks	0	0	0	0
Satisfied with Life	0	0	0	+
Basic Activities of Daily Living				
Eating				
Did[c]	-	-	0	0
Could[d]	-	0	0	0
Transferring				
Did	-	0	0	0
Could	0	0	0	0
Ambulating				
Did	0	0	0	-
Could	0	0	0	0
Bathing				
Did	0	+	0	0
Could	0	+	0	0
Instrumental Activities of Daily Living				
Medications				
Did Take	0	0	0	0
Could Take	+	0	0	0

SOURCE: Four- and eight-month surveys

NOTE: *p*-Values are based on standard errors inflated to account for the effects of clustering and weighting.

+ = estimated treatment group value exceeds estimated control group value, $p \leq 0.10$.
- = estimated treatment group value falls below estimated control group value, $p \leq 0.10$.
0 = estimated difference between treatment and control group values with $p > 0.10$.

[a] "Equal weight" refers to the analysis in which observations are weighted to give agencies equal representation.

[b] "Size weight" refers to the analysis in which observations are weighted to represent agencies proportional to their size.

[c] "Did" refers to items on the survey asking whether patients usually performed the activity.

[d] "Could" refers to items asking patients whether they *could* have performed the activity had there been no one to help. If we had measured only whether a patient actually performed an activity, the comparison between treatment and control agencies might have been distorted because control agencies were more likely to have been providing services at four months, and patients receiving relatively more assistance might appear to have been more impaired because they had fewer opportunities to perform activities without help.

TABLE IV.10

SUMMARY OF DIFFERENT WEIGHTING SCHEMES ON ESTIMATED EFFECTS OF PROSPECTIVE PAYMENT ON OUTCOMES FROM THE QUALITY ASSURANCE AND CLAIMS DATA

	Equal Weight[a]	Size Weight[b]
Quality Assurance Data		
Grooming		
Improvement	0	0
Stabilization	0	0
Bathing		
Improvement	0	0
Stabilization	0	0
Toileting		
Improvement	0	0
Stabilization	0	0
Transferring		
Improvement	0	0
Stabilization	0	0
Ambulating		
Improvement	0	0
Stabilization	0	0
Light Meal Preparation		
Improvement	0	0
Stabilization	0	0
Housekeeping		
Improvement	0	0
Stabilization	0	0
Management of Medications		
Improvement	0	0
Stabilization	0	0
Pain		
Improvement	0	0
Stabilization	0	0
Most Problematic Ulcer		
Improvement	0	0
Stabilization	0	0
Surgical Wound Status		
Improvement	–	–
Stabilization	0	0
Dyspnea		
Improvement	0	0
Stabilization	0	0
Urinary Incontinence or Catheter Present		
Improvement	0	+
Stabilization	+	0
Confusion		
Improvement	+	0
Stabilization	0	0

TABLE IV.10 *(continued)*

	Equal Weight[a]	Size Weight[b]
Behavior Problem Frequency		
Improvement	0	0
Stabilization	0	0
Emergency Visits		
Hospital emergency room	0	-
Outpatient clinic or urgent care clinic	-	-
Physician's office	-	-
Any of the above	-	-
Medicare Claims Data		
Mortality		
Within 120 days	0	0
Within eight months	0	0
Within one year	0	-
Admission to Hospital for Same-Body-System Diagnosis		
Within 120 days	0	0
Within eight months	-	-
Within one year	-	-
Admission to SNF for Same-Body-System Diagnosis		
Within 120 days	0	-
Within eight months	0	-
Within one year	0	0
Admission to HHA for Same-Body-System Diagnosis		
Within 120 days	+	+
Within eight months	+	+
Within one year	+	+

SOURCE: Quality assurance and Medicare claims data

NOTE: p-Values are based on standard errors inflated to account for the effects of clustering and weighting.

SNF = skilled nursing facility; HHA = home health agency.

+ = estimated treatment group value exceeds estimated control group value, $p \leq 0.10$.
- = estimated treatment group value falls below estimated control group value, $p \leq 0.10$.
0 = estimated difference between treatment and control group values with $p > 0.10$.

[a] "Equal weight" refers to the analysis in which observations are weighted to give agencies equal representation.

[b] "Size weight" refers to the analysis in which observations are weighted to represent agencies proportional to their size.

the discrepancies (see Appendix B). The negative difference on improvement in surgical wound status remained robust to the two weighting schemes, and we did not identify any obvious outlying small or large agencies to explain the negative treatment effect. As we discussed in Section B.3, although we cannot confirm or rule out this negative surgical wound impact with certainty, the context of an absence of other impacts suggests that it probably does not represent a real finding. The main feature of Tables IV.9 and IV.10 is the absence of any pattern or trend in treatment-control differences across the numerous outcomes.

The apparent demonstration effect of increased home health admission was explained by an extreme outlier agency in the treatment group, VNA of Dade County, Florida. This agency had a mean value for the outcome nearly double that of the next highest agency (see Appendix B). From our site visits, we learned that the Health Care Financing Administration mistakenly reported in mid-1997 that it was closing the agency for violations of Medicare's Conditions of Participation. The erroneous report, in conjunction with a highly competitive home health care market, caused the agency to lose large numbers of patients to other agencies. Exclusion of this agency from the agency equal analysis essentially eliminated the differences in home health admission.

Several other differences observed in the agency equal analysis did hold up in the agency size analysis. As shown in Table IV.10, reductions in emergency care use and hospitalization for same-body-system diagnosis were robust to the agency size analysis. (In Section C.1, we have discussed our reasons for accepting these results with caution.) Treatment-control differences in the three satisfaction outcomes of staff rushed through work, staff encouraged independence, and staff paid attention to patient also remained in the agency size analysis (Table IV.11). Agency-level scatterplots of these outcomes did not reveal any obviously extremely small or large agencies that

TABLE IV.11

SUMMARY OF DIFFERENT WEIGHTING SCHEMES ON ESTIMATED EFFECTS OF PROSPECTIVE PAYMENT ON SATISFACTION OUTCOMES FROM THE FOUR-MONTH SURVEY

	Equal Weight[a]	Size Weight[b]
General Satisfaction with Agency		
Percentage Extremely Satisfied	0	0
Would Recommend	0	0
Discharged Too Soon	0	0
Needed Services--Not Available and a Big Problem	0	0
Satisfied with Personal Care Arrangements		
Four months	0	-
Eight months	0	0
Satisfaction with Agency Staff		
Staff Arrived Three Hours Late or More or Failed to Come at All[c]		
Little or none of the time	0	0
Staff Rushed Through Work[c]		
Little or none of the time	0	0
Most or all of the time	+	+
Staff Encouraged Independence[d]		
Strongly agree	0	0
Disagree or strongly disagree	+	+
Staff's Reassurance and Emotional Support[e]		
Excellent	0	0
Staff Paid Attention to Patient[e]		
All of the time	0	0
Some of the time, or little or none of the time	+	+
Aide		
Aide Completed All Work	0	0
Aide Did Not Come Often Enough	0	0
Nurses and Therapists		
Nurses and Therapists Did Not Come Often Enough	0	0
Nurses' and Therapists' Care and Thoroughness in Examination and Treatment[e]		
Excellent	0	0
Nurses' and Therapists' Visits Long Enough[c]		
All of the time	0	0

TABLE IV.11 *(continued)*

	Equal Weight[a]	Size Weight[b]
Nurses and Therapists Gave Clear Explanations of Medical Conditions and Treatment[c]		
All of the time	0	0
Nurses' and Therapists' Teaching About Care[e]		
Excellent	0	0

SOURCE: Four-month patient survey.

NOTE: *p*-Values are based on standard errors inflated to account for the effects of clustering and weighting.

+ = estimated treatment group value exceeds estimated control group value, $p \leq 0.10$.
- = estimated treatment group value falls below estimated control group value, $p \leq 0.10$.
0 = estimated difference between treatment and control group values with $p > 0.10$.

[a]"Equal weight" refers to the analysis in which observations are weighted to give agencies equal representation.

[b]"Size weight" refers to the analysis in which observations are weighted to represent agencies proportional to their size.

[c]Response categories were: all of the time, most of the time, some of the time, and little or none of the time.

[d]Response categories were: strongly agree, agree, disagree, and strongly disagree.

[e]Response categories were: excellent, very good, good, fair, and poor.

could be influencing the results, suggesting the treatment-control differences were broad based and more likely represented true impacts (data not shown).

2. Including Patients with More than One Demonstration Home Health Admission

Adding back observations we had excluded because they were not "first admissions" (that is, they were patients who had previously had admissions to demonstration agencies) did not affect the results or our conclusions. Including these patients in the analyses did not substantially alter any of the estimated treatment-control differences or change the significance levels. We therefore conclude that our results are the same regardless of whether we restrict the sample to first admissions.

F. SUMMARY

We do not detect any adverse impacts from prospective payment on a wide array of patient health and function outcomes. If anything, we find weak evidence that this payment method was associated with a slight *decline* in hospitalization for same-body-system diagnoses (two percentage point reduction at one year; 7 percent relative to the control group mean) and in use of emergency care during the at-risk period (two percentage point reduction; 14 percent relative to the control group mean).

Our analyses do provide evidence of patient dissatisfaction in three measures of interpersonal care at four months after home health admission: (1) staff rushing through work, (2) staff encouraging patients' independence, and (3) staff paying attention to patients. These impacts were small in absolute size (roughly two to four percentage points in magnitude). However, given the small number of dissatisfied patients, the impacts were substantial (47 to 64 percent relative to the control group mean).

The main goals of home health care are to maintain or restore health and function, and to prevent medical complications. Therefore, the absence of impacts on patients' health and functioning is a positive finding for the prospective payment system in this demonstration, especially in light of the marked reductions in service provision that occurred. Problems with patient dissatisfaction should be more easily correctable than problems with decrements in function or symptoms.

V. SUBGROUP IMPACTS

If policymakers wish to consider appropriate policy responses to prospective payment, they must be aware of circumstances under which this system leads to differential impacts on patient outcomes. Patient outcomes may vary across agency or patient subgroups because the incentives of prospective payment may cause certain types of agencies to behave differently or certain types of patients to be treated differently. For example, for-profit agencies may respond more eagerly than nonprofit agencies to the profit incentives of the demonstration, or agencies may reduce home health visits and episode durations more for patients who have access to informal caregivers than for patients who lack these caregivers. An understanding of subgroup effects also helps us to project demonstration results to a nationwide prospective payment system, even if demonstration agencies do not mirror the national distribution of agencies.

To investigate whether prospective payment affected the quality of care differentially for subgroups, we estimated impacts on key patient outcomes within selected agency and patient subgroups. We conducted statistical tests to determine (1) whether the program impact within a given subgroup was significantly different from zero; and (2) whether a significant difference existed in the impacts between subgroup pairs (for example, between for-profit and nonprofit agencies).[1]

A. DEFINING AGENCY AND PATIENT SUBGROUPS

To define our subgroups, we selected agency and patient characteristics most likely to modify the way that agencies altered their service provision under prospective payment. Agencies respond behaviorally to prospective payment primarily by making changes in service provision, and effects

[1]As discussed in Section A.1 of Chapter III, we do not present any agency-level or patient-level subgroup analyses of survey data outcomes because of problems relating to the small sample sizes.

on patient outcomes would be mediated through those changes. The five pairs of agency subgroups that we examined were:

1. ***"High-Use" or "Low-Use" Predemonstration Practice Pattern.*** We defined high-use agencies as those with a base-year case-mix–adjusted index of service provision above the median, and low-use agencies as those with an index equal to or below the median.

2. ***For Profit or Nonprofit***

3. ***Small or Large Size.*** We defined small agencies as those providing fewer than 30,000 visits during the base year and defined large agencies as those providing 30,000 visits or more.

4. ***Hospital-Based or Freestanding***

5. ***Below or Above Cost Limits.*** We defined agencies below cost limits as those whose base-year costs per visit were at or below the Section 223 limits and defined agencies above cost limits as those with base-year costs above the Section 223 limits.

The three patient characteristics that we examined were:[2]

1. ***Independent or Not Independent in Taking Oral Medications***

2. ***Other Caregiver Available or Not Available.*** We defined patients to have other care available if they received assistance from paid help or another person residing in the home, or if they resided in an assisted-living facility.

3. ***High or Low Expected Costs.*** We defined a patient to have high expected costs if he or she fell in the top quartile of predicted costs for all patients in our sample.[3]

---

[2]We obtained all the information used to form these patient characteristic subgroups from the quality assurance (QA) forms that the agencies completed at admission. As discussed in Chapter II, these data were available for only a subsample of patients, and the precision of our estimates was weaker than it would have been had the full sample been available.

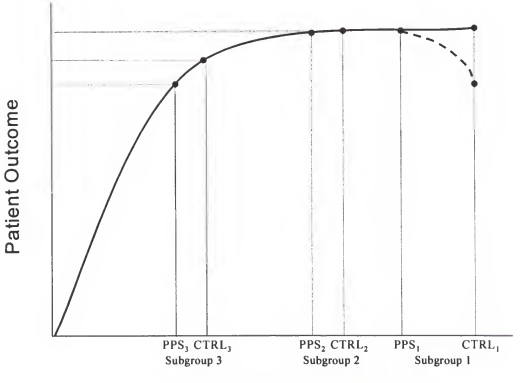
[3]The variables used to calculate predicted costs are those identified by Phillips et al. (1992) as increasing costs for the average patient by more than 30 percent.

B. POTENTIAL SUBGROUP EFFECTS ON PATIENT OUTCOMES

Most likely, impacts of reductions in the amount of home health care on patient outcome depend both on the amount of home health care provided at baseline (the control group mean) and on the magnitude of the reduction in the amount of home health care. The solid curve in Figure V.1 shows a hypothesized relationship between a patient outcome and the amount of home health care, and the effects of reductions in the amount of home health care in three agency subgroups. In agency subgroup one, prospective payment causes a relatively large reduction in the amount of home health care levels, but the patient outcome remains unaffected. Agency subgroup two experiences a smaller reduction than subgroup one as a result of the demonstration and, also as in subgroup one, there is no impact on patient outcome. For both subgroup one and subgroup two, the curve is flat (that is, the control agencies in both subgroups are providing optional services that are not necessary to achieve the health outcome). In contrast, agency subgroup three reduces home health care by the same amount as does subgroup two, but the patient outcome is adversely affected, because subgroup three is at the level of service at which more services do produce better patient outcomes (that is, the curve is steep in that region). As discussed in Chapter IV, Section A, it is also possible that the patient outcome-home health care volume curve may start to descend at high volumes of home health care (shown as a dashed segment of the curve in Figure V.1), in which case a decline in the amount of home health care in subgroup 1 might even lead to an improvement in patient outcome. In Chapter IV, Section A we also raised the possibility that prospective payment might lead agencies to actually shift the curve upwards and to the left (not shown in Figure V.1, see Figure IV.1), in which case a decline in services might also be associated with an improvement in patient outcome for certain subgroups.

FIGURE V.1

CURVE OF HYPOTHESIZED RELATIONSHIP BETWEEN PATIENT OUTCOME AND AMOUNT OF HOME HEALTH CARE SERVICE AND POSSIBLE SUBGROUP EFFECTS



NOTE: The solid curve shows a hypothesized relationship between a patient outcome and the amount of home health care provided. In subgroup 1, prospective payment causes the amount of home health care delivered to fall, but no change in patient outcome occurs. In subgroup 2, prospective payment causes a smaller decrease than in subgroup 1 in the amount of home health care delivered, but again, no change in patient outcome occurs. In subgroup 3, prospective payment causes a decrease in the amount of home health care delivered of roughly the same size as in subgroup 2, but here, patient outcome does decrease. The dashed segment shows the possibility of a descending portion of the curve, in which case a fall in the amount of home health care in subgroup 1 actually leads to an improvement in patient outcome. Prospective payment may also lead to a upwards and leftwards shift in the curve (not shown), in which case a decline in amount of home health care could also be associated with an improvement in patient outcome for certain subgroups.

PPS = prospectively paid agency; CTRL = control group agency.

In the patient subgroups, which define needy, vulnerable patients, our concern is that home health service reductions will manifest as adverse impacts on health and functional outcomes, even though they do not do so in "typical" patients. We expect that, compared to "typical" patients, the outcome-service curves for these more impaired patients are shifted to the right. In other words, for much of the curve, more services produce better outcomes (the curve ascends), and the point at which more services produce no additional benefits (the curve flattens) occurs only at high levels of service. The same comments made above on ascending, flat, and (possible) descending portions of the curve, and on possible upwards and leftward movements of the curve still apply, however.

C. IMPACTS FOR AGENCY SUBGROUPS

The results of the agency-level subgroup analyses presented here are consistent with our findings from our analysis of main demonstration effects. In the majority of outcomes for which we find no overall demonstration effect, we also find no effect within subgroups. In the case of same-body-system hospitalizations, for which there was a possible demonstration effect, we find the same effect within each subgroup. The consistency of the overall analysis and subgroup analyses suggests that the demonstration operated uniformly across different types of agencies.

The general absence of agency-level subgroup impacts on patient outcomes contrasts with the evidence of agency-level subgroup impacts on home health service provision found by Trenholm (2000). We summarize the impacts on the provision of home health services here:

- High-use agencies made larger (but proportionally similar relative to the control group means) cuts in the number and duration of services than did low-use agencies.

- For-profit agencies and nonprofit agencies made cuts in service that were not statistically different.[4]

- Small agencies made larger cuts (both absolutely, and proportionally relative to the control group means) in the number and duration of services than did large agencies, but these results were of borderline statistical significance.

- Freestanding agencies made larger (but proportionally similar relative to the control group means) cuts in the number and duration of services than did hospital-based agencies, but these results were not statistically significant.

- Agencies above the cost limit made larger cuts (both absolutely and proportionally relative to the control group means) in the number and duration of services than did agencies below the limit.

1. High-Use or Low-Use Agencies

We conclude that prospective payment did not lead to impacts on quality of care within either the high-use agency subgroup or the low-use agency subgroup. We also conclude that it did not lead to differential quality impacts between this pair of subgroups (Table V.1).

a. Differences Within Subgroups

It is likely that the mixed treatment-control differences within subgroups are statistical artifacts. Within the high-use agency subgroup, two treatment-control differences favored treatment agencies (a small effect on stabilization in dyspnea and a moderate effect on improvement in confusion), and two differences favored control agencies (a small effect on stabilization in toileting and a small effect on stabilization in transferring). The observed reductions in the number of visits provided by the treatment agencies in the high-use subgroup--a nearly 30 percent reduction in nursing and aide visits

[4]However, nonprofit agencies made larger reductions in physical therapy visits. The difference may have been the result of the relatively large therapy staffs of many nonprofit agencies, which afforded the agencies a greater opportunity to make large-scale reductions in services.

TABLE V.1

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY HIGH-USE OR LOW-USE PRACTICE PATTERN
(Percentages, Unless Otherwise Noted)

Outcome	High-Use Practice Pattern			Low-Use Practice Pattern			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect)
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	21.5	−0.8 (0.47)	−3.7	20.6	−1.4 (0.24)	−6.8	0.6 (0.68)
At Eight Months	29.7	−1.5 (0.23)	−5.1	28.8	−2.6* (0.09)	−9.0	1.1 (0.56)
At One Year	35.9	−2.5* (0.07)	−7.0	34.2	−2.4 (0.17)	−7.0	−0.1 (1.0)
Medical Symptoms							
Pain							
Improvement	50.9	3.7 (0.20)	7.3	54.3	0.8 (0.73)	1.5	2.9 (0.42)
Stabilization	83.7	1.7 (0.31)	2.0	82.4	1.1 (0.31)	1.3	0.6 (0.76)
Dyspnea							
Improvement	42.6	4.0 (0.25)	9.4	49.3	0.6 (0.83)	1.2	3.4 (0.42)
Stabilization	78.8	3.5* (0.09)	4.4	79.8	−0.1 (0.91)	−0.1	3.6 (0.12)
Confusion							
Improvement	34.2	9.6*** (0.01)	28.1	43.1	2.6 (0.53)	6.0	7.0 (0.19)
Stabilization	84.5	1.4 (0.59)	1.7	85.4	2.8** (0.05)	3.3	1.4 (0.52)

TABLE V.1 *(continued)*

Outcome	High-Use Practice Pattern			Low-Use Practice Pattern			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	52.3	-3.0 (0.33)	-5.7	56.2	-0.8 (0.77)	-1.4	-2.2 (0.55)
Stabilization	87.2	-2.2 (0.13)	-2.5	87.6	2.0 (0.24)	2.3	-4.2** (0.04)
Bathing							
Improvement	50.3	-0.8 (0.79)	-1.6	53.6	-1.3 (0.64)	-2.4	0.5 (0.90)
Stabilization	84.9	-1.2 (0.44)	-1.4	83.8	1.9 (0.19)	2.3	-3.1* (0.10)
Toileting							
Improvement	54.4	1.4 (0.62)	2.6	56.2	-3.2 (0.30)	-5.7	4.6 (0.23)
Stabilization	91.1	-1.7* (0.10)	-1.9	91.1	2.1 (0.18)	2.3	-3.8** (0.04)
Transferring							
Improvement	43.9	1.3 (0.70)	3.0	51.3	-4.8 (0.14)	-9.4	6.1 (0.15)
Stabilization	88.1	-2.4* (0.07)	-2.7	88.1	0.5 (0.89)	0.6	2.9* (0.10)
Ambulating							
Improvement	30.6	0.5 (0.88)	1.6	33.8	-1.8 (0.52)	-5.3	2.3 (0.55)
Stabilization	91.2	-1.5 (0.14)	-1.6	88.7	0.3 (0.87)	0.3	1.8 (0.18)
Instrumental ADLs							
Management of Oral Medications							
Improvement	35.9	-2.8 (0.41)	-7.8	38.0	-1.8 (0.60)	-4.7	1.0 (0.82)
Stabilization	87.5	-0.8 (0.61)	-0.9	86.6	2.0 (0.15)	2.3	2.8 (0.17)

TABLE V.1 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

[c] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

and a 9 percent reduction in therapy visits, both relative to the control group mean--are unlikely to have led to this precise, specific pattern of effects. Within the low-use agency subgroup, we find only a small treatment-control difference in stabilization in confusion favoring the treatment group, again incompatible with the relative reductions by the low-use treatment agencies of 26 percent in nursing visits and of 17 percent in aide visits.

b. Differences Across Subgroups

The small differences between high-use and low-use subgroups in treatment effects for stabilization in the basic activities of daily living (ADLs) of grooming, bathing, toileting, and transferring are also unlikely to represent true impacts. For each of these four outcomes, we find small, negative effects of prospective payment in the high-use subgroup (ranging from 1.4 to 2.7 percent relative to the control mean), and small, positive effects in the low-use subgroup (ranging from 0.6 to 2.3 percent relative to the control mean). However, the estimated effects for improvement in the ADLs of toileting, transferring, and ambulating are in the opposite direction, with positive effects of prospective payment in the high-use subgroup and negative effects in the low-use subgroup (though lacking significant *p*-values for subgroup effects). It is unclear how the substantial reductions in the number of visits by treatment agencies in the high-use subgroup would simultaneously increase the likelihood of improvement in basic ADLs while reducing the likelihood of stabilization. Moreover, we find no adverse effects of prospective payment among high-use agencies for other outcomes (same-body-system hospitalizations, medical symptoms, and instrumental activities of daily living). Even though treatment agencies in the high-use group made larger cuts in the number of visits than did low-use treatment agencies, treatment agencies in the high-use subgroup were still providing more visits during the year after admission than were treatment agencies in the low-use subgroup (68 and 39 visits, respectively). By definition, high-use

treatment agencies had been providing a higher intensity of services than had low-use control agencies in the year before the demonstration started and could thus likely make greater reductions without affecting outcomes. Finally, the negative effects in the high-use subgroup on the four stabilization outcomes are too small to be a major policy concern.[5]

2. For-Profit or Nonprofit Agencies

In response to prospective payment, both for-profit agencies and nonprofit agencies made equally large reductions in the quantity and duration of home health services, and we find no real demonstration effects in outcomes of medical symptoms and functioning either within or between for-profit and nonprofit agency subgroups (Table V.2). There appear to be statistically significant positive treatment effects on improvement in confusion and on stabilization in confusion in the for-profit subgroup. In addition, we find a small differential impact of borderline significance between the subgroups in stabilization in oral medication management favoring treatment agencies in the for-profit subgroup and favoring control agencies in the nonprofit subgroup. Neither the effects on confusion nor the effect on stabilization in oral medication management are likely to be real. Because we are conducting many statistical tests, these results are probably merely "false positives" (that is, they are statistically significant only by chance). The lack of a plausible mechanism that would explain these isolated results strengthens this conclusion.

---

[5]The large number of statistical tests also greatly increased the likelihood of a Type II error, or false-positive result, occurring. Moreover, the stabilization outcomes were more vulnerable than the improvement outcomes to Type II errors because the samples for the stabilization outcomes were larger, and the stabilization outcome control group means were higher (80 to 90 percent versus roughly 50 percent for the improvement outcomes). Statistical power for binary outcomes is greater at extreme mean values (for example, 10 percent or 90 percent) than at mean values of roughly 50 percent.

Nevertheless, we find what appear to be differences between the for-profit and nonprofit agencies in same-body-system hospitalizations (Table V.2). Substantial reductions in these hospitalizations by for-profit agencies were apparent by 120 days and continued to grow at the eight-month and one-year time points. By one year, for-profit treatment agencies had five percentage points fewer same-body-system hospitalizations than did for-profit control agencies (13 percent of the control group mean). In contrast, there was no treatment-control difference in the nonprofit agencies at one year.

Despite the magnitude of the observed reductions, we do not believe that this subgroup difference represents a true impact. First, this finding conflicts with the findings of the other analyses of the for-profit/nonprofit subgroup (that is, of no impacts). Second, both for-profit and nonprofit prospectively paid agencies made equally substantial reductions in home health services. Thus, for-profit agencies would have had to implement an innovative strategy, which agency staff did not mention in our site visits, that led to a sizeable, isolated reduction in same-body-system hospitalizations without affecting on any other health or functional outcomes.

We considered whether an unmeasured difference between treatment and control agencies, such as an imbalance in control agencies' intrinsic hospitalization patterns, could explain the observed results. To explore this possibility, we constructed an agency-level base-quarter hospitalization rate variable (defined in Section C.1.a. of Chapter IV) to examine the sensitivity of our results.[6] We first ran a regression on the full sample with the base-quarter hospitalization rate variable as a new control

[6]One of the main classification variables of the Home Health Utilization Group's base-quarter case-mix adjuster is whether a patient has an intervening or "embedded" hospital stay during the 120 days after a home health admission. It is therefore easy to calculate the percentage of each agency's patients hospitalized during the 120-day period in the base-quarter. This variable may partially capture otherwise unmeasurable local practice patterns of physicians or hospitals that refer to the agency.

TABLE V.2

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY FOR-PROFIT OR NONPROFIT STATUS
(Percentages, Unless Otherwise Noted)

Outcome	For-Profit Status			Nonprofit Status			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	22.2	-3.9*** (0.00)	-17.6	20.1	1.6 (0.25)	8.0	-5.5*** (0.01)
At Eight Months	30.4	-4.8*** (0.00)	-15.8	28.4	0.4 (0.79)	1.4	-5.2** (0.04)
At One Year	36.2	-4.8*** (0.01)	-13.3	34.2	-0.1 (0.94)	-0.3	4.7* (0.06)
Medical Symptoms							
Pain							
Improvement	50.7	1.6 (0.53)	3.2	54.0	2.9 (0.33)	5.4	1.3 (0.72)
Stabilization	83.2	0.1 (0.97)	0.1	83.1	2.7* (0.08)	3.2	-2.6 (0.16)
Dyspnea							
Improvement	43.2	2.1 (0.48)	4.9	47.6	2.4 (0.51)	5.0	0.3 (0.96)
Stabilization	78.2	1.4 (0.50)	1.8	80.4	1.9 (0.38)	2.4	-0.5 (0.89)
Confusion							
Improvement	34.6	9.0** (0.04)	26.0	41.3	3.2 (0.42)	7.7	5.8 (0.35)
Stabilization	82.3	3.7** (0.03)	4.5	87.7	0.5 (0.87)	0.6	3.2 (0.21)

TABLE V.2 *(continued)*

Outcome	For-Profit Status			Nonprofit Status			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	51.1	-3.1 (0.28)	-6.1	57.1	-0.7 (0.83)	-1.2	-2.4 (0.58)
Stabilization	85.5	0.6 (0.80)	0.7	89.4	-1.1 (0.46)	-1.2	1.7 (0.50)
Bathing							
Improvement	48.3	-0.5 (0.87)	-1.0	55.3	-1.5 (0.62)	-2.7	1.0 (0.83)
Stabilization	83.3	1.7 (0.29)	2.0	85.7	-1.4 (0.43)	-1.6	3.1 (0.17)
Toileting							
Improvement	52.4	-0.9 (0.78)	-1.7	58.1	-0.8 (0.80)	-1.4	-0.1 (0.98)
Stabilization	89.6	0.9 (0.47)	1.0	92.7	-1.7 (0.22)	-1.8	2.6 (0.18)
Transferring							
Improvement	42.7	-0.6 (0.87)	-1.4	51.3	-2.9 (0.43)	-5.7	2.3 (0.63)
Stabilization	86.3	-0.8 (0.48)	-0.9	90.0	-1.3 (0.31)	-1.4	0.5 (0.81)
Ambulating							
Improvement	30.5	-0.7 (0.83)	2.3	33.3	-0.7 (0.82)	-2.1	0.0 (0.99)
Stabilization	89.7	0.1 (0.99)	0.1	90.7	-1.6 (0.14)	-1.8	1.7 (0.34)
Instrumental ADLs							
Management of Oral Medications							
Improvement	34.8	-0.3 (0.94)	-0.9	38.9	-4.2 (0.24)	-10.8	3.9 (0.45)
Stabilization	86.0	2.2 (0.16)	2.6	88.5	-1.4 (0.42)	-1.6	3.6* (0.10)

TABLE V.2 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

[c] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

variable. We then ran a second regression using the original model (without the new control variable) on a smaller sample from which we had dropped all cases from the four agencies with the highest base-quarter hospitalization rates. All four agencies were for-profit control agencies in Illinois. In both regressions, the treatment impacts in the for-profit subgroup were attenuated considerably, and, in the second regression, the for-profit subgroup effect ceased to be statistically significant (the regression results are shown in Appendix B). The attenuation of the subgroup impacts by this base-quarter hospitalization rate variable bolsters the argument that additional unmeasured characteristics of for-profit agencies, unrelated to prospective payment, contributed to the observed subgroup effect.

3. Small or Large Agencies

Small agencies reduced the number of visits by substantially more than did large agencies both in absolute terms and relative to the control group means in each subgroup, although these results were of borderline statistical significance. However, we observe no consistent pattern of differences to suggest true impacts on patient functional status outcomes in agencies of either subgroup (Table V.3). In small agencies, a large effect on improvement in pain favors the treatment group (9.2 percentage points; 22.4 percent of the control group mean). However, with the exception of borderline effects on stabilization in pain and stabilization in confusion that also favor the treatment group, we find no other treatment-control differences among the small agencies. In large agencies, the only significant treatment-control difference is a large effect on improvement in confusion that favored the treatment group (7.9 percentage points; 20.6 percent of the control group mean). Once again, we cite a lack of plausibility and the high risk of Type II error to conclude that these results do not reflect substantive findings. Reflecting the overall findings, we find reductions in the number of same-body-system hospitalizations of comparable size among both large agencies and small

TABLE V.3

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY SMALL OR LARGE SIZE
(Percentages, Unless Otherwise Noted)

Outcome	Large Size			Small Size			Estimated Difference of Subgroup Differences[b] (p-Value for Difference in Subgroup Effect)
	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	21.1	-1.2 (0.18)	-5.7	21.6	-0.8 (0.73)	-3.7	-0.4 (0.86)
At Eight Months	29.2	-2.0* (0.06)	-6.8	30.2	-2.5 (0.36)	-8.3	0.5 (0.86)
At One Year	35.0	-2.0* (0.09)	-5.7	36.1	-3.6 (0.19)	-10.0	1.6 (0.59)
Medical Symptoms							
Pain							
Improvement	54.2	0.2 (0.92)	0.4	41.0	9.2** (0.04)	22.4	9.0* (0.06)
Stabilization	83.6	0.8 (0.51)	1.0	80.6	3.4* (0.09)	4.2	2.6 (0.22)
Dyspnea							
Improvement	47.1	1.1 (0.62)	2.3	35.4	6.4 (0.34)	18.1	-5.3 (0.44)
Stabilization	80.2	1.0 (0.44)	1.2	73.3	3.8 (0.37)	5.2	-2.8 (0.54)
Confusion							
Improvement	38.4	7.9*** (0.01)	20.6	34.7	0.5 (0.93)	1.4	7.4 (0.29)
Stabilization	86.3	1.1 (0.43)	1.3	76.5	5.2* (0.10)	6.8	4.1 (0.25)

TABLE V.3 *(continued)*

Outcome	Large Size			Small Size			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect)
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	54.3	-1.1 (0.61)	-2.0	51.4	-4.5 (0.41)	-8.8	3.4 (0.55)
Stabilization	88.1	-1.3 (0.26)	-1.5	82.9	3.0 (0.33)	3.6	-4.3 (0.17)
Bathing							
Improvement	52.9	-1.0 (0.67)	-1.9	44.2	-1.2 (0.84)	-2.7	0.2 (0.98)
Stabilization	84.9	0.3 (0.89)	0.4	81.7	0.3 (0.96)	0.4	0.0 (1.0)
Toileting							
Improvement	55.5	-1.2 (0.64)	-2.2	52.6	0.1 (0.98)	0.2	-1.3 (0.82)
Stabilization	91.5	-0.6 (0.45)	-0.7	88.8	0.8 (0.75)	0.9	-1.4 (0.54)
Transferring							
Improvement	47.6	-1.8 (0.47)	-3.8	42.4	-1.9 (0.76)	-4.5	0.1 (0.99)
Stabilization	88.9	-2.3* (0.07)	-2.6	83.2	2.9 (0.26)	3.5	-5.2* (0.06)
Ambulating							
Improvement	31.8	0.6 (0.78)	1.9	32.1	-5.1 (0.37)	-15.9	5.7 (0.33)
Stabilization	90.6	-0.8 (0.36)	-0.9	87.7	-0.6 (0.76)	-0.7	-0.2 (0.88)
Instrumental ADLs							
Management of Oral Medications							
Improvement	37.0	-1.3 (0.60)	-3.5	35.2	-5.7 (0.36)	-16.2	4.4 (0.46)
Stabilization	87.5	0.7 (0.68)	0.8	85.3	-0.2 (0.89)	-0.2	0.9 (0.73)

TABLE V.3 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

[c] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

agencies, although the effects do not reach statistical significance in the smaller small-agency subgroup.

4. Hospital-Based or Freestanding Agencies

In our report on home health use, we cautiously concluded that freestanding agencies made larger reductions in the number of visits than did hospital-based agencies (Trenholm 2000).[7] However, we find no corresponding subgroup impacts on patient outcomes in the analyses presented here (Table V.4). In both hospital-based agencies and freestanding agencies, prospective payment essentially had no differential effects on hospital admission for same-body-system diagnosis, medical symptoms, or ADLs. The treatment-control differences in same-body-system hospitalizations reached statistical significance only in the freestanding agency subgroup, but the magnitudes of the estimates were similar in the hospital-based agencies (where there were many fewer observations and, therefore, less statistical power).

5. Below or Above Cost Limits Agencies

Finally, although agencies above the cost limits reduced services significantly more than did agencies below the limits (by roughly twice as much), prospective payment had negligible impacts on patient functional outcomes in agencies, whether under or over their cost limits (Table V.5).[8]

---

[7]First, freestanding agencies made large and significant cuts in all visit types (nursing, aide, and therapy), whereas hospital-based agencies made significant cuts in only one type (nursing). Second, despite the lack of precision for the hospital-based subgroup effect, the reduction in home health aide visits by freestanding agencies was significant; in contrast, the change that hospital-based agencies made was insignificant.

[8]The larger treatment impact in visit reductions in agencies above the limits was partly due to control agencies in this subgroup having substantially increased their visits per patient between the year before the demonstration started and the demonstration's first year (Trenholm 2000). Among agencies below the limits, neither treatment nor control agencies had made such changes prior to the start of the demonstration.

TABLE V.4

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY WHETHER HOSPITAL BASED OR FREESTANDING
(Percentages, Unless Otherwise Noted)

Outcome	Hospital Based			Freestanding			Estimated Difference of Subgroup Differences[a] (p-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (p-Value[c])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[c])	Estimated Difference as Percentage of Control Group Mean	
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	20.5	-3.7 (0.12)	-18.0	21.3	-0.7 (0.34)	-3.3	-3.0 (0.22)
At Eight Months	28.2	-2.9 (0.32)	-10.3	29.6	-2.0** (0.04)	-6.8	-0.9 (0.75)
At One Year	33.8	-2.7 (0.41)	-8.0	35.5	-2.4** (0.03)	-6.8	-0.3 (0.91)
Medical Symptoms							
Pain							
Improvement	51.9	4.6 (0.30)	8.9	52.3	1.9 (0.37)	3.6	2.7 (0.58)
Stabilization	85.4	-0.8 (0.67)	-0.9	82.7	1.7 (0.14)	2.1	-2.5 (0.28)
Dyspnea							
Improvement	41.1	8.0 (0.17)	19.5	46.1	1.4 (0.56)	3.0	6.6 (0.28)
Stabilization	82.4	-2.6 (0.52)	-3.2	78.6	2.3* (0.10)	2.9	-4.9 (0.26)
Confusion							
Improvement	36.1	9.6 (0.17)	26.6	38.1	5.6** (0.05)	14.7	4.0 (0.58)
Stabilization	89.4	1.6 (0.68)	1.8	84.0	2.2 (0.18)	2.6	0.6 (0.93)

TABLE V.4 *(continued)*

Outcome	Hospital Based			Freestanding			Estimated Difference of Subgroup Differences[b] (p-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	52.9	−0.5 (0.94)	−0.9	54.1	−2.1 (0.33)	−3.9	1.6 (0.80)
Stabilization	90.8	0.0 (0.95)	0.0	86.7	−0.2 (0.72)	−0.2	0.2 (0.93)
Bathing							
Improvement	50.9	3.8 (0.52)	7.5	51.8	−1.7 (0.45)	−3.3	5.5 (0.34)
Stabilization	88.2	1.7 (0.64)	1.9	83.7	0.1 (0.98)	0.1	1.6 (0.64)
Toileting							
Improvement	52.9	4.0 (0.53)	7.6	55.5	−1.6 (0.50)	−2.9	5.6 (0.37)
Stabilization	93.6	0.2 (0.97)	0.2	90.6	−0.2 (0.59)	−0.2	0.4 (0.85)
Transferring							
Improvement	48.0	−2.7 (0.73)	−5.6	46.6	−1.6 (0.51)	−3.4	−1.1 (0.89)
Stabilization	91.4	−1.9 (0.51)	−2.1	87.4	−0.9 (0.31)	−1.0	−1.0 (0.75)
Ambulating							
Improvement	31.3	−0.8 (0.90)	−2.6	32.0	−0.7 (0.75)	−2.2	−0.1 (0.98)
Stabilization	92.9	−0.3 (0.86)	−0.3	89.7	−0.8 (0.30)	−0.9	0.5 (0.85)
Instrumental ADLs							
Management of Oral Medications							
Improvement	33.9	0.7 (0.92)	2.1	37.3	−2.7 (0.29)	−7.2	3.4 (0.64)
Stabilization	89.0	1.0 (0.75)	1.1	86.8	0.4 (0.81)	0.5	0.6 (0.83)

TABLE V.4 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a]This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b]This value is the difference of the estimated treatment-control differences for each subgroup.

[c]This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.

TABLE V.5

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY WHETHER UNDER OR OVER (BASE-YEAR) COST LIMITS
(Percentages, Unless Otherwise Noted)

Outcome	Under Cost Limit			Over Cost Limit			Estimated Difference of Subgroup Differences[b] (p-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	21.3	-1.8** (0.05)	-8.5	20.5	2.8 (0.20)	13.7	-4.6* (0.06)
At Eight Months	29.4	-2.3** (0.04)	-7.8	29.0	-1.1 (0.66)	-3.8	-1.2 (0.67)
At One Year	35.3	-2.6** (0.03)	-7.4	34.8	-1.4 (0.62)	-4.0	-1.2 (0.72)
Medical Symptoms							
Pain							
Improvement	51.4	3.0 (0.14)	5.8	56.6	-1.3 (0.75)	-2.3	4.3 (0.32)
Stabilization	83.5	0.6 (0.62)	0.7	81.4	5.5*** (0.00)	6.8	4.9*** (0.00)
Dyspnea							
Improvement	43.6	2.1 (0.42)	4.8	54.1	3.4 (0.48)	6.3	-1.3 (0.82)
Stabilization	79.2	1.3 (0.40)	1.6	79.3	3.6 (0.17)	4.5	-2.3 (0.40)
Confusion							
Improvement	36.4	6.1** (0.04)	16.8	44.8	6.1 (0.35)	13.6	0.0 (0.99)
Stabilization	84.9	1.7 (0.27)	2.0	84.7	4.4 (0.14)	5.2	2.7 (0.30)

TABLE V.5 *(continued)*

Outcome	Under Cost Limit			Over Cost Limit			Estimated Difference of Subgroup Differences[b] (p-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	53.5	-1.9 (0.44)	-3.6	56.2	-2.0 (0.62)	-3.6	0.1 (0.98)
Stabilization	87.4	0.0 (0.90)	0.0	87.2	-1.5 (0.48)	-1.7	1.5 (0.56)
Bathing							
Improvement	51.0	-0.3 (0.91)	-0.6	55.2	-5.0 (0.24)	-9.1	4.7 (0.34)
Stabilization	84.4	0.7 (0.66)	0.8	84.9	-1.8 (0.36)	-2.1	2.5 (0.26)
Toileting							
Improvement	54.8	-1.1 (0.67)	-2.0	56.7	0.4 (0.94)	0.7	1.5 (0.79)
Stabilization	91.1	0.2 (0.96)	0.2	90.8	-2.4 (0.19)	-2.6	2.6 (0.27)
Transferring							
Improvement	46.1	-1.4 (0.63)	-3.0	51.2	-3.7 (0.50)	-7.2	2.3 (0.72)
Stabilization	88.2	-1.4 (0.18)	-1.6	87.3	0.8 (0.81)	0.9	2.2 (0.35)
Ambulating							
Improvement	31.2	-0.7 (0.77)	-2.2	35.5	-0.7 (0.87)	-2.0	0.0 (1.0)
Stabilization	90.6	-0.9 (0.29)	-1.0	87.9	0.2 (0.98)	0.2	1.1 (0.55)
Instrumental ADLs							
Management of Oral Medications							
Improvement	36.1	-1.5 (0.59)	-4.2	39.8	-6.0 (0.21)	-15.1	4.5 (0.41)
Stabilization	87.6	0.3 (0.85)	0.3	85.0	1.4 (0.60)	1.6	1.1 (0.66)

TABLE V.5 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

[b] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

Both subgroups showed treatment effects of reduced same-body-system hospitalizations that were not substantially different between subgroups in either magnitude or direction. (The difference is significant only in the subgroup of agencies that were below the cost limits.) In addition, as in the other subgroup analyses, we find some subgroup differences, such as stabilization in pain, that achieved statistical significance, but that were small in size and not meaningful in the absence of any other findings.

D. IMPACTS FOR PATIENT SUBGROUPS

Agencies made significantly larger reductions in home health services for the high-expected-cost patient subgroup than for the low-expected-cost subgroup (although relative to the control group mean, the reduction was of a similar proportion; Trenholm 2000). In this analysis, we find no corresponding difference in patient outcomes between these two patient subgroups and, therefore, no evidence to suggest that prospectively paid agencies had overzealously reduced care for high-cost patients (Table V.6). Not surprisingly, patients with high expected costs had somewhat lower control group means for the outcome measures because they were somewhat less likely to improve or stabilize. However, the estimated treatment-control differences between the two subgroups were small and comparable.

Finally, in the home health use report, there were no subgroup differences in service provision for patients who could and who could not take oral medications or for patients with informal caregivers and for those without caregivers (Trenholm 2000). We also find no subgroup differences in patient outcomes in either of these subgroups (Tables V.7 and V.8, respectively). There is no recognizable pattern in the hospitalization and medical symptom outcomes. There is a relatively large positive treatment effect on improvement in confusion in the subgroup of patients who could

TABLE V.6

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY HIGH OR LOW EXPECTED PATIENT COSTS
(Percentages, Unless Otherwise Noted)

	High Expected Costs			Low Expected Costs			
Outcome	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Estimated Difference of Subgroup Differences[b] (p-Value for Difference in Subgroup Effect)[c]
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	22.4	-0.6 (0.72)	-2.7	19.7	-0.4 (0.72)	-2.0	-0.2 (0.96)
At Eight Months	31.0	-1.2 (0.55)	-3.9	27.6	-0.4 (0.81)	-1.4	-0.8 (0.65)
At One Year	37.1	-1.4 (0.49)	-3.8	32.7	-0.1 (0.97)	-0.3	-1.3 (0.47)
Medical Symptoms							
Pain							
Improvement	49.8	4.0* (0.10)	8.0	53.6	1.3 (0.57)	2.4	2.7 (0.25)
Stabilization	81.6	1.2 (0.36)	1.5	83.9	1.1 (0.32)	1.3	0.1 (0.92)
Dyspnea							
Improvement	44.8	1.5 (0.60)	3.3	45.5	1.3 (0.58)	2.9	0.2 (0.92)
Stabilization	78.0	1.7 (0.32)	2.2	79.8	1.9 (0.18)	2.4	-0.2 (0.81)
Confusion							
Improvement	34.0	5.5* (0.08)	16.2	40.1	5.7* (0.07)	14.2	-0.2 (0.96)
Stabilization	81.0	2.0 (0.28)	2.5	86.8	2.1 (0.12)	2.4	-0.1 (0.73)

TABLE V.6 *(continued)*

Outcome	High Expected Costs			Low Expected Costs			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect)[z]
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	51.7	-2.8 (0.28)	-5.4	55.8	-1.8 (0.43)	-3.2	-1.0 (0.57)
Stabilization	85.0	-1.1 (0.49)	-1.3	88.5	0.3 (0.80)	0.3	-1.4 (0.26)
Bathing							
Improvement	48.8	-0.6 (0.85)	-1.2	53.5	-1.3 (0.59)	-2.4	0.7 (0.74)
Stabilization	83.7	0.4 (0.81)	0.5	84.8	0.5 (0.65)	0.6	-0.1 (0.93)
Toileting							
Improvement	53.7	-0.5 (0.87)	-0.9	56.4	-2.0 (0.46)	-3.5	1.5 (0.53)
Stabilization	88.1	0.6 (0.59)	0.7	92.6	-0.5 (0.63)	-0.5	1.1 (0.29)
Transferring							
Improvement	44.1	-1.2 (0.65)	-2.7	48.9	-2.2 (0.40)	-4.5	1.0 (0.61)
Stabilization	86.2	-1.2 (0.45)	-1.4	89.0	-1.0 (0.39)	-1.1	-0.2 (0.94)
Ambulating							
Improvement	30.2	-1.0 (0.68)	-3.3	32.9	-0.9 (0.69)	-2.7	-0.1 (0.92)
Stabilization	90.3	-1.1 (0.34)	-1.2	90.1	-0.2 (0.87)	-0.2	-0.9 (0.33)
Instrumental ADLs							
Management of Oral Medications							
Improvement	32.5	-1.3 (0.66)	-4.0	39.6	-2.9 (0.29)	-7.3	1.6 (0.50)
Stabilization	85.5	0.3 (0.87)	0.4	87.9	0.9 (0.48)	1.0	0.6 (0.64)

TABLE V.6 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

[c] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

* Significantly different from zero at the 0.10 level, two-tailed test.

TABLE V.7

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY WHETHER PATIENT CAN TAKE ORAL MEDICATIONS INDEPENDENTLY
(Percentages, Unless Otherwise Noted)

Outcome	Can Take Oral Medications Independently			Cannot Take Oral Medications Independently			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect[c])
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	18.5	-0.8 (0.58)	-4.3	22.0	-0.1 (0.96)	-0.5	-0.7 (0.66)
At Eight Months	26.6	-1.2 (0.49)	-4.5	30.3	-0.1 (0.97)	-0.3	-1.1 (0.55)
At One Year	31.7	-1.2 (0.51)	-3.8	36.1	-0.1 (0.97)	-0.3	-1.1 (0.58)
Medical Symptoms							
Pain							
Improvement	55.9	0.9 (0.66)	1.6	48.8	3.2 (0.16)	6.6	-2.3 (0.24)
Stabilization	84.6	0.6 (0.62)	0.7	82.0	1.5 (0.19)	1.8	-0.9 (0.46)
Dyspnea							
Improvement	46.9	1.5 (0.53)	3.2	44.0	1.3 (0.60)	3.0	0.2 (0.90)
Stabilization	81.1	1.6 (0.23)	2.0	77.6	2.1 (0.21)	2.7	-0.5 (0.85)
Confusion							
Improvement	47.0	9.5*** (0.01)	20.2	35.5	4.9* (0.10)	13.8	4.6 (0.12)
Stabilization	89.3	2.6** (0.05)	2.9	80.4	1.7 (0.33)	2.1	0.9 (0.25)

TABLE V.7 *(continued)*

Outcome	Can Take Oral Medications Independently			Cannot Take Oral Medications Independently			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect)
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	70.2	-1.1	-1.6	46.9	-2.5	-5.3	1.4
		(0.68)			(0.29)		(0.57)
Stabilization	90.9	0.3	0.3	84.0	-0.4	-0.5	0.7
		(0.83)			(0.76)		(0.59)
Bathing							
Improvement	62.2	-0.8	-1.3	44.9	-1.1	-2.4	0.3
		(0.75)			(0.69)		(0.84)
Stabilization	87.1	0.9	1.0	82.0	0.2	0.2	0.7
		(0.46)			(0.91)		(0.48)
Toileting							
Improvement	76.0	-3.8	-5.0	50.3	-1.0	-2.0	-2.8
		(0.22)			(0.70)		(0.31)
Stabilization	94.9	0.2	0.2	87.6	-0.1	-0.1	0.3
		(0.86)			(0.93)		(0.78)
Transferring							
Improvement	55.0	-0.1	-0.2	42.8	-2.7	-6.3	2.6
		(0.97)			(0.26)		(0.23)
Stabilization	90.4	-0.7	-0.8	86.1	-1.3	-1.5	0.6
		(0.59)			(0.29)		(0.62)
Ambulating							
Improvement	33.5	-0.7	-2.1	30.7	-1.0	-3.3	0.3
		(0.78)			(0.65)		(0.81)
Stabilization	91.9	-0.4	-0.4	88.7	-0.5	-0.6	0.1
		(0.70)			(0.66)		(0.96)

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

TABLE V.7 *(continued)*

[c] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

TABLE V.8

ESTIMATED DIFFERENCES BETWEEN PROSPECTIVELY PAID AGENCIES AND COST-REIMBURSED AGENCIES IN SELECTED PATIENT OUTCOMES, BY WHETHER THE PATIENT HAS OTHER CAREGIVER
(Percentages, Unless Otherwise Noted)

	No Other Caregiver			Other Caregiver			
Outcome	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (p-Value[a])	Estimated Difference as Percentage of Control Group Mean	Estimated Difference of Subgroup Differences[b] (p-Value for Difference in Subgroup Effect[c])
Hospital Admission for Same-Body-System Diagnosis							
At 120 Days	19.2	0.0 (0.99)	0.0	21.2	-0.8 (0.57)	-3.8	0.8 (0.60)
At Eight Months	28.0	-1.7 (0.31)	-6.1	29.1	0.0 (0.99)	0.0	-1.7 (0.38)
At One Year	34.4	2.5 (0.14)	-7.3	34.0	0.5 (0.79)	1.5	2.0 (0.15)
Medical Symptoms							
Pain							
Improvement	52.9	2.2 (0.33)	4.2	51.7	2.1 (0.32)	4.1	0.1 (0.94)
Stabilization	83.4	1.2 (0.33)	1.4	83.0	1.1 (0.31)	1.3	0.1 (0.95)
Dyspnea							
Improvement	45.6	0.0 (0.99)	0.0	45.0	2.5 (0.31)	5.6	-2.5 (0.14)
Stabilization	79.9	1.0 (0.53)	1.3	78.7	2.4* (0.09)	3.0	1.4 (0.19)
Confusion							
Improvement	38.7	7.5** (0.02)	19.4	37.2	4.7 (0.12)	12.6	2.8 (0.22)
Stabilization	85.6	2.2 (0.14)	2.6	84.3	2.1 (0.15)	2.5	0.1 (0.88)

TABLE V.8 *(continued)*

Outcome	No Other Caregiver			Other Caregiver			Estimated Difference of Subgroup Differences[b] (*p*-Value for Difference in Subgroup Effect)
	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
Basic ADLs							
Grooming							
Improvement	62.7	-3.5 (0.23)	-5.6	49.3	-1.4 (0.51)	-2.8	-2.1 (0.30)
Stabilization	89.6	-0.6 (0.68)	-0.7	85.6	0.2 (0.89)	0.2	-0.8 (0.43)
Bathing							
Improvement	56.4	-0.6 (0.83)	-1.1	48.7	-1.2 (0.63)	-2.5	0.6 (0.78)
Stabilization	85.7	0.3 (0.80)	0.4	83.5	0.7 (0.63)	0.8	-0.4 (0.78)
Toileting							
Improvement	64.8	-4.7 (0.12)	-7.3	51.4	-0.1 (0.97)	-0.2	-4.6** (0.06)
Stabilization	93.6	-1.0 (0.36)	-1.1	89.1	0.5 (0.60)	0.6	-1.5 (0.14)
Transferring							
Improvement	50.6	0.2 (0.93)	0.4	44.9	-2.8 (0.28)	-6.2	3.0* (0.08)
Stabilization	89.1	-1.5 (0.26)	-1.7	87.3	-0.7 (0.53)	-0.8	-0.8 (0.41)
Ambulating							
Improvement	31.3	1.0 (0.70)	3.2	32.2	-2.0 (0.36)	-6.2	3.0** (0.03)
Stabilization	91.2	-1.5 (0.15)	-1.6	89.5	0.3 (0.81)	0.3	-1.8* (0.06)
Instrumental ADLs							
Management of Oral Medications							
Improvement	42.5	-1.1 (0.68)	-2.6	33.9	2.9 (0.29)	-8.6	1.8 (0.37)
Stabilization	89.5	-0.2 (0.87)	-0.2	85.1	1.4 (0.32)	1.6	1.6* (0.08)

TABLE V.8 *(continued)*

SOURCE: Data on hospital admission for same-body-system diagnosis are from the Medicare claims data. The remaining outcomes are from the quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies. Observations have been weighted to give agencies equal representation in the analysis. Estimated differences have been adjusted through logit models to control for preexisting differences between treatment and control agencies.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This value is the difference of the estimated treatment-control differences for each subgroup.

[c] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

ADL = activities of daily living.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.

take oral medications independently, and a smaller positive treatment effect on improvement in confusion in the subgroup of patients who could not take these medications independently, both of which are statistically significant. The subgroup difference for this outcome is not significant. It does not seem plausible that prospectively paid agencies would perform exceptionally well only on this single outcome, so we consider it another statistical artifact.

E. SUMMARY

The absence of demonstration effects across subgroups on patient outcomes contrasts with the substantial subgroup differences in service reduction (Trenholm 2000). In the agency subgroup analysis, high-use, small, freestanding, and above-the-limit agencies reduced services more deeply than did their subgroup counterparts, yet we find no evidence of corresponding impacts on patient outcomes. In the patient subgroup analysis, high-expected-cost patients were subject to larger service reductions, but without a detectable difference in patient outcomes.

The lack of correlation between the magnitude of the service reductions and patient outcomes suggests that, in each subgroup, agencies were operating well within the area of the service-outcome curve at which additional services are provided that do not produce improvements in patient outcomes; that is, where the curve is flat. The subgroups we studied are more like subgroups one and two in Figure V.1, which reduced services to varying degrees but without producing impacts on patient outcomes in any case, than like subgroup three, which reduced services and worsened patient outcomes.

The subgroup analyses in this chapter thus confirm the main findings of the preceding chapter--that prospective payment had no discernible impacts on patient outcomes. Prospective payment as implemented in this demonstration appears to preserve the quality of care across different agency

and patient subgroups, providing additional evidence that a program of this type could be implemented without affecting patient health outcomes.

VI. CONCLUSIONS

The Per-Episode Home Health Prospective Payment Demonstration substantially reduced the number and duration of services provided to patients, but we do not detect any impacts on a broad range of measures of patients' health, symptoms, and functioning. We find weak evidence that prospective payment may have reduced emergency visits to clinics and physicians' offices, as well as hospitalizations for diagnoses related to the home health care diagnosis. There also is some evidence that, for a small number of patients, prospective payment increased patient dissatisfaction with specific aspects of interpersonal care. Our findings provide additional evidence that prospective payment holds great potential as a payment reform for Medicare home health care.

A. SUMMARY OF RESULTS

We consistently find throughout our analysis that prospective payment had negligible effects on the quality of care. The perceptions of treatment agency patients and control agency patients of overall health and the two groups' self-reports of days confined to bed did not differ. We also find little evidence to support any impacts of prospective payment on basic and instrumental activities of daily living. We conclude that prospective payment had no impact on several medical symptoms and outcomes (such as pain, dyspnea, and mortality). Furthermore, the demonstration had no observed impacts on admission to skilled nursing facilities (SNFs) or home health agencies for same-body-system diagnoses. An apparent increase in same-body-system admissions to home health agencies for treatment agency patients resulted from a fluke that caused a single treatment group agency in Florida to suffer an acute loss of its patients to competing agencies.

We do find weak evidence that the number of emergency visits to clinics and physicians' offices and hospitalizations for same-body-system diagnoses fell among treatment agencies. During the at-

risk period, there was a two percentage point reduction (15 percent relative to the control group mean) in emergency visits of any type (to a hospital emergency room, physician's office, or clinic). This reduction resulted from drops in the number of visits to hospital outpatient clinics (one percentage point; 33 percent relative to the control group mean) and to physicians' offices (0.5 percentage point; 38 percent relative to the control group mean). The reduction in hospitalizations for same-body-system diagnoses appeared by 120 days and continued to widen over time. By one year, treatment agency patients had a 2.4 percentage point lower incidence of same-body-system hospitalizations (seven percent relative to the control group mean).

The reductions in the two types of health services use were somewhat surprising. We considered three possible explanations: (1) the impacts are real and result from concurrent cuts in home health services and improvements in the quality of care by the treatment agencies; (2) the impacts are real and result from cuts in home health services without effects on the quality of care, as additional home health visits generate increased hospital and emergency care use that has no marginal benefit; and (3) the differences result from random baseline differences in unmeasured characteristics of treatment and control agencies. We further explored the latter possibility to the limited extent that our data allowed and found that measures of predemonstration agency hospitalization patterns partially explained, but did not eliminate, the observed differences. We thus cautiously conclude that the impacts may be real; however, our data do not enable us to distinguish the first explanation from the second. We suspect that the second explanation is correct, because we have no other evidence that treatment agencies actually improved quality, and the study by Weinberger et al. (1996) supports the idea that increased provision of medical services might induce even more service use.

Prospective payment did not affect patients' satisfaction with either the overall care their agency provided or the technical quality of care nurses and therapists provided. In contrast, we find increased dissatisfaction among patients of prospectively paid agencies with three facets of interpersonal care by agency staff: (1) rushing through work, (2) failing to encourage patients' independence, and (3) failing to pay attention to what patients had to say. More treatment agency patients complained that agency staff rushed through work most or all of the time (by a difference of two percentage points; 50 percent relative to the control group mean). In addition, more treatment agency patients complained that agency staff did not encourage them to be independent (by a difference of four percentage points; 50 percent relative to the control group mean). Finally, more treatment agency patients complained that agency staff did not pay much attention to them (by a difference of three percentage points; 60 percent relative to the control group mean). These impacts, although large relative to the control group mean, affected only a small number of patients and were thus small in absolute terms. Even though other data suggest that treatment agencies' visits were at least as long as those of control agencies, we wonder whether treatment agency staff may have tried to perform more work during visits. Because they made fewer visits overall (relative to control agency staff), they may therefore have appeared more rushed and less responsive.

Consistent with the main analysis, the analysis of patient outcomes in agency and patient subgroups revealed no differential impacts on patient outcomes. The absence of impacts was notable in view of the substantial agency and patient subgroup differences in the way in which prospective payment reduced home health services.

B. LIMITATIONS

The seeming lack of results of our study (that is, no major effects of prospective payment on health and functional outcomes) raises the question of whether we could have missed any important

impacts on the quality of care. In Table III.1, we presented minimum detectable differences for our different data sources under different assumptions about the intercorrelation of responses of patients within each agency. The table shows that, even under the assumption of high intraagency correlation ($k = 0.1$), we should be able to detect a roughly 10 percentage point treatment-control difference in most outcomes. For the majority of outcomes, the intraagency correlation was in the range of $k = 0.01$, so we should be able to detect differences of four to five percentage points. Although it is difficult to determine what minimum differences are meaningful for the outcomes under study, differences in the range of four to five percentage points should be relevant to policymakers, especially in light of the substantial service reductions observed in the demonstration.

We measured a wide range of outcomes, but the list is not exhaustive. It is possible that we failed to measure outcomes that might be more sensitive to decreases in the number of home health visits. However, these outcomes, whatever they might be, probably affect only a small percentage of the Medicare home health population.

The generalizability of the study's results may be limited. For example, the study cannot determine what will happen to the quality of care and to patient outcomes if a national program leads to even larger cuts in service than those observed in the demonstration. Under a national program, agencies would presumably not be protected from financial losses, possibly prompting some to respond with even deeper reductions. Moreover, per-episode payments to agencies would probably be based on regional or national averages, rather than on the agencies' own historical costs per episode. It is likely that agencies with high-use practice patterns would then face losses. To reduce their costs per episode sufficiently to remain financially solvent, these high-use agencies might have to make deeper service cuts than those in the demonstration, possibly leading to adverse effects.

The voluntary nature of the demonstration raises the possibility that participating agencies may not have been representative of home health agencies nationwide. However, we find that agency characteristics that might be expected to influence agency behavior under prospective payment (predemonstration practice patterns, size, auspice, profit status, and whether above or below cost limits) did not affect our results of a general absence of impacts on quality of care outcomes. Variables for agency characteristics were unimportant in the main treatment-control analyses, and per-episode prospective payment did not have differential impacts for different types of agency in the analyses of agency subgroups defined by agency characteristics. That agency characteristics did not affect our findings suggests that our results can be widely generalized.

Furthermore, demonstration agencies (especially the prospectively paid ones), knowing they were being scrutinized as subjects in a demonstration, may have made greater attempts to maintain the quality of care than they otherwise would have done. In this case, we may have underestimated the impacts of prospective payment on the quality of care. However, agencies indicated to us during our site visits that, other than learning the new demonstration quality assurance (QA) process, they did not change their QA procedures during the demonstration.

C. POLICY IMPLICATIONS

This report demonstrates that prospective payment can accomplish the policy goal of substantially decreasing home health care services without harming patients. This information will be useful to HCFA as it prepares to implement the congressionally mandated Medicare home health prospective payment system.

One of the main advantages of a prospective payment system over other options for constraining home health care expenditures is that it gives agencies the flexibility (subject to the payment limits and quality oversight) to deliver care as they see fit. In contrast, regulatory approaches, such as fixed

expenditure caps or rigid algorithms of covered benefits, do a poor job of accommodating the clinical complexities that each beneficiary's case presents. This limitation is especially true for home health, as providers must contend with the home and family situation as well as with the health status of patients. This demonstration shows that, given such flexibility, agencies were able to safely reduce home health service use.

Although patients' increased level of dissatisfaction with care under prospective payment is a less serious issue than are any adverse impacts on health and functioning, it does warrant some attention. The impacts we observe in this demonstration are quite small (two percentage points), but one could imagine they could become greater in a nationwide system. HCFA may wish to consider monitoring patient satisfaction under a prospective payment system for home health care.

Finally, as policymakers are well aware, although this report demonstrates that home health services *can* be significantly reduced without negatively affecting the quality of care, it obviously does not *guarantee* that quality of care will be maintained. HCFA has thus devoted extensive attention and resources to developing a workable system for ongoing QA. It is appropriate that HCFA plans to regularly and carefully monitor the quality of care as prospective payment is implemented.

REFERENCES

Brown, Randall, Barbara Phillips, Christine Bishop, Amy Klein, Grant Ritter, Craig Thornton, Peter Schochet, and Kathleen Skwara. "The Effects of Predetermined Payment Rates for Home Health Care." Princeton, NJ: Mathematica Policy Research, Inc., December 1995.

Cheh, Valerie. Princeton, NJ: Mathematica Policy Research, Inc., forthcoming

Cheh, Valerie and Christopher Trenholm. "Preliminary Report: The Impact of Prospective Payment on the Cost per episode: Striking the Balance Between Decreasing Use and Increasing Cost." Draft report, Princeton, NJ: Mathematica Policy Research, Inc., February 1998.

Chen, Arnold, and Helaine Noveck. "Preliminary Report: The Impact of Prospective Payment on Medicare Home Health Quality of Care." Princeton, NJ: Mathematica Policy Research, Inc., January 1998.

Cox, Brenda G., and Steven B. Cohen. *Methodological Issues for Health Care Surveys.* New York: Marcel Dekker, Inc., 1985.

Fisher, Elliott S., and H. Gilbert Welch. "Avoiding the Unintended Consequences of Growth in Medical Care: How Might More Be Worse?" *Journal of the American Medical Association,* vol. 281, no. 5, February 3, 1999, pp. 446-453.

Health Care Financing Administration, Office of Strategic Planning. *Health Care Financing Review. Medicare and Medicaid Statistical Supplement, 1999.* Baltimore, MD: HCFA, OSP.

Health Care Financing Administration, Office of Strategic Planning. *Health Care Financing Review. Medicare and Medicaid Statistical Supplement, 1996.* Baltimore, MD: HCFA, OSP.

Hosmer, David W., and Stanley Lemeshow. *Applied Logistic Regression.* New York: John Wiley and Sons, Inc., 1989.

Hughes, Susan L., Alec Ulasevich, Frances M. Weaver, William Henderson, Larry Manheim, Joseph D. Kubal, and Frank Bonarigo. "Impact of Home Care on Hospital Days: A Meta Analysis." *Health Services Research*, vol. 32, no. 4, October 1997, pp. 415-432.

Kemper, Peter. "The Evaluation of the National Long Term Care Demonstration: Overview of the Findings." *Health Services Research*, vol. 23, no. 1, April 1988, pp.162-174.

Leon, Joel, Patricia Neuman, and Stephen Parente. "Understanding the Growth in Medicare's Home Health Expenditures." Report submitted to the Henry J. Kaiser Family Foundation. Bethesda, MD: Project HOPE Center for Health Affairs, June 1997.

Phillips, Barbara. "A Pilot Study of the Adequacy of Post-Hospital Community Care for the Elderly." Report submitted to the Health Care Financing Administration. Princeton, NJ: Mathematica Policy Research, Inc., January 1990.

Phillips, Barbara R., Randall S. Brown, Jennifer L. Schore, Amy C. Klein, Peter Z. Schochet, Jerrold W. Hill, and Dexter O. Chu. "Case-Mix Analysis Using Demonstration Data: Home Health Prospective Payment Demonstration." Draft report. Princeton, NJ: Mathematica Policy Research, Inc., December 1992.

Phillips, Barbara R., Randall Brown, Jennifer Schore, Amy C. Klein, Peter Z. Schochet, Jerrold W. Hill, and Craig Thornton. "Evaluation of the Per-Episode Home Health Prospective Payment Demonstration Design." Report submitted to the Health Care Financing Administration. Princeton, NJ: Mathematica Policy Research, Inc., September 1995.

Phillips, Barbara, and Rachel Thompson. "Transition Within a Turbulent System: An Analysis of the Initial Implementation of the Per-Episode Home Health Prospective Payment Demonstration." Draft report. Princeton, NJ: Mathematica Policy Research, Inc., February 1997.

PriceWaterhouseCoopers LLP. "Congress Changes Medicare Reimbursement for Home Health Services: What Does It Mean for Home Health Providers." *Health Policy Series*, no. 6, September 1998.

Prospective Payment Assessment Commission. *Report and Recommendations to Congress.* Washington, DC: ProPAC, March 1, 1996.

Rich, Michael W., Valerie Beckham, Carol Wittenberg, Charles L. Leven, Kenneth E. Freedland, and Robert M. Carney. "A Multidisciplinary Intervention to Prevent the Readmission of Elderly Patients with Congestive Heart Failure." *New England Journal of Medicine*, vol. 333, no. 18, November 2, 1995, pp. 1190-1195.

Schore, Jennifer. "Regional Variation in the Use of Medicare Home Health Services." In *Persons with Disabilities: Issues in Health Care Financing and Service Delivery*, edited by Joshua M. Wiener, Steven B. Clauser, and David L. Kennell. Washington, DC: The Brookings Institution, 1995.

Schore, Jennifer. "Patient, Agency, and Area Characteristics Associated with Regional Variation in the Use of Medicare Home Health Services." Report submitted to the Health Care Financing Administration. Princeton, NJ: Mathematica Policy Research, Inc., September 1994.

Schore, Jennifer. "The Impact of Home Health Prospective Payment on Medicare Service Use and Reimbursement." Report submitted to Health Care Financing Administration. Princeton, NJ: Mathematica Policy Research, Inc., March 2000.

Shaughnessy, Peter W., Robert E. Schlenker, and David F. Hittle. "Home Health Care Outcomes Under Capitated and Fee-for-Service Payment." *Health Care Financing Review,* vol. 16, no. 1, fall 1994, pp.187-222.

Shaughnessy, Peter W., Kathryn S. Crisler, Robert E. Schlenker, and Angela G. Arnold. "Outcome-Based Quality Improvement in Home Care." *Caring*, vol. 14, no. 2, February 1995, pp. 44-49.

Teplitsky, Sanford, V., and Mary Ann Janson (eds.). *Home Health & Hospice Manual: Regulations and Guidelines*. Owings Mills, MD: National Health Publishing, 1985-1992.

Trenholm, Christopher. "Per-Episode Prospective Payment for Medicare Home Health Care Sharply Reduces Service Use." Draft report submitted to the Health Care Financing Administration. Princeton, NJ: Mathematica Policy Research, Inc., May 2000.

U.S. General Accounting Office. *Medicare Home Health Benefit: Impact of Interim Payment System and Agency Closures on Access to Services*. Washington, DC: GAO, September 1998.

Weinberger, Morris, Eugene Z. Oddone, and William G. Henderson. "Does Increased Access to Primary Care Reduce Hospital Admissions: Veterans Affairs Cooperative Study Group on Primary Care and Hospital Readmission." *New England Journal of Medicine,* vol. 334, no. 22, May 1996, pp. 1441-1447.

Welch, H. Gilbert, David E. Wennberg, and W. Pete Welch. "The Use of Medicare Home Health Services." *New England Journal of Medicine,* vol. 335, no. 5, August 1, 1996, pp. 324-329.

APPENDIX A

DESIGN AND FIELDING OF THE PATIENT SURVEY

In this appendix, we describe the design and fielding of the patient survey. Key design issues concerned the timing of the interviews relative to home health admission and the design of the sample to support analyses with the data sometimes weighted to represent agencies equally and sometimes weighted to represent patients equally. The patient survey was fielded in two waves of computer-assisted telephone interviews.

A. MAJOR DESIGN ISSUES FOR THE PATIENT SURVEY

The timing of the administration of the patient interviews was designed to capture both the short- and long-run impacts of per-episode prospective payment. To support analysis of the effects of the lump-sum payment, the four-month interview was administered shortly before the 120th day after home health admission. This survey provided "snapshot" measures of the use of non-Medicare services relatively close to the time of discharge for most patients (sometime during the first 90 days), as well as measures covering the entire four-month interval since home health admission.

The eight-month interview was designed to capture any longer-run effects of the lump-sum payment, as well as the effect of per-visit payments during the outlier period. When payment shifted from the lump-sum payment to outlier payment, agencies may have had an incentive to *increase* the number of visits rendered to patients still in their care (provided that per-visit rates were adequate to cover the marginal cost of the visits). In fact, however, agencies did not act on this incentive; they continued to reduce visits during the outlier period (Trenholm 1999). Of course, when we designed the patient survey, we did not know that they would do so.

A key design feature of the patient survey is that both waves of interviews were administered a specified number of days after *admission*, rather than at discharge or at a specified number of days after discharge. In this way, the data for both the treatment and control groups were collected at a comparable point in a spell of illness. Thus, differences in the timing of data collection cannot be

confounded with the effects of the payment methodology. Given that the average duration of home health stays decreased substantially under per-episode payment, this comparability of the timing for the treatment and control data is a major methodological advantage of the survey data.

B. FIELDING OF THE PATIENT SURVEY

Sample intake for the four-month survey began with patients admitted in January 1997 and ended with patients admitted in December 1997. The initiation of sample intake varied by agency, with intake beginning first for the agencies enrolling earliest in the demonstration.

For most agencies, sample intake continued for six months. This six-month period fell during the middle or latter part of the agencies' second demonstration year. It occurred during the middle of the second demonstration year for the last agencies to enroll, and during the latter part of the second demonstration year for the first agencies to enroll. Intake was extended to seven or eight months for the smallest agencies to help ensure that the minimum number of complete interviews was obtained. In a few cases, sampled admissions for the smallest agencies occurred early in the third year of their demonstration operations.

Administration of the telephone interviews began in early May 1997, approximately four months after sample intake began. Administration of the four-month interview continued through April 1998, and administration of the eight-month interview through August 1998.

We obtained a completion rate of more than 90 percent for both the four-month and eight-month interviews. The completion rate is the ratio of the number of completed interviews to the number of eligible cases. We learned after some interviews had been completed that the cases were ineligible. Ineligible cases are discussed in detail in Section B.

Most home health admissions during the sample intake period were identified from Palmetto Government Benefits Administration (PGBA) claims files, although some were identified by the

agencies. (See Section B for a discussion of the construction of the universe from which the patient sample was drawn.) After we received claims files--about the middle of each month--we added admissions from lists submitted by the agencies and drew the sample. We then mailed advance letters to the sampled patients and asked the agencies to provide contact information (including each patient's telephone number and name of next of kin) via fax or courier, to facilitate reaching sampled patients quickly.

We usually were able to complete the four-month interview by the 120th day, as planned. Overall, 87 percent were completed by that time. Our inability to complete a minority (13 percent) of the four-month interviews until after the 120th day was not a problem, as the agencies did not alter their behavior after the at-risk period, despite a possible incentive to do so, and there was essentially no difference between the treatment and control groups in the proportion of interviews completed by the 120th day. Most of the four-month interviews that had not been completed by the 120th day were completed during the subsequent two weeks. In all, 96 percent of the interviews were completed by the 134th day after home health admission. (The timing of the administration of the eight-month interview was not critical.)

C. THE PATIENT SURVEY SAMPLE

1. Universe

The universe for the patient survey consisted of patients newly admitted to demonstration home health agencies for whom an agency was eligible to receive a lump-sum payment (in the case of treatment agencies) or would have been eligible under demonstration rules (in the case of control agencies).[1] The largest group of newly admitted patients for whom an agency was *ineligible* for

[1]Three types of cases were to be excluded because agencies were not paid a lump sum for them. (continued...)

lump-sum payments consisted of those who had been discharged and were then readmitted either (1) fewer than 165 days after the initial admission, or (2) before a 45-day gap in home health services had occurred.

a. Identifying Patients in the Universe

In the case of patients admitted during the *second* half of the month, the time lags between admission and claims filing and processing necessitated that we draw the sample from claims *three* months before the fielding date. We were able to complete all the tasks required in preparation for fielding. For patients admitted during the first half of the month, however, we did not have time to complete the tasks and begin the interview on schedule. Nevertheless, we have no reason to believe that patients admitted during the latter part of a given month were systematically different from patients admitted earlier in that month.

Overall, we were able to identify 78 percent of expected admissions during the second half of a month from claims in process and those for which processing was complete. However, drawing admissions only from the second half of the month was a problem for two groups of agencies because the number of patients admitted during that period was insufficient to produce the desired number of admissions. One group consisted of very small agencies for which it was necessary to select all, or almost all, admissions during the entire month in order to obtain the desired number. The other group consisted of a few larger agencies for which claims processing appeared to have lagged.

---

[1](...continued)
These cases are those in which: 1) the patient was enrolled in a managed care organization, 2) Medicare was a secondary payer, and 3) no home health visits were provided.

We developed a separate procedure for identifying the universe for these two groups of agencies. We asked the staff of the very small agencies to provide us with lists of patients admitted each month for the duration of the sample intake period. We asked the staff of the larger agencies to continue to give us these lists until we could identify a sufficient proportion of their admissions from claims. All the agencies in the two groups agreed to the requests. The monthly lists of patient admissions were data entered, and the resulting files were combined with the claims data files. Duplicate entries were eliminated before the sample was drawn.

b. Limitations of the Patient-Identification Approach

The procedures we implemented were designed to include as many newly admitted patients as possible, but the drawback is that we include observations on patients who were not eligible for a lump-sum payment and therefore not appropriate for inclusion in the study. The inclusion of these ineligible observations is the chief limitation of this approach. (As we describe in Section B.4, we subsequently dropped ineligible cases from the patient survey analysis files). Six months after an admission, say, it would have been possible to use claims to accurately identify the universe of patients eligible for a lump-sum payment. By that time, claims would have been relatively complete. Therefore, it would have been reasonable to assume that a gap in claims represented a break in service, and to apply the demonstration rules defining an episode. Moreover, after six months, the agencies would have corrected most claims submitted in error (for example, for managed care enrollees or for patients for whom Medicare was a secondary payer). Nevertheless, because we had to avoid delays in identifying the sample, a nontrivial proportion of claims were still outstanding. Consequently, we could not rely on claims to identify the universe accurately.

2. Review of Sample Design

We designed the patient survey sample to permit weighting of observations to represent agencies equally, or to represent them according to their size (measured in demonstration patient episodes), without introducing extremely large weights for some observations under either weighting scheme. Introducing these weights would have seriously reduced the precision of our estimates.

To reach our target of 2,000 completed eight-month interviews of eligible patients, we estimated that we had to complete 2,500 four-month interviews with these patients, and to draw a sample of 2,980 cases (Brown et al. 1995). A sample draw of this size was necessary because some patients selected for the patient survey would be identified later as ineligible for demonstration services. It also allowed for the usual sources of attrition, including death occurring between the four- and eight-month interviews, inability to locate the sampled patient, and refusal to participate.[2]

3. Implementation of Sample Selection

For the survey sample, we selected a set fraction of demonstration admissions to each agency still participating in the demonstration during the second demonstration year, provided that the fraction was expected to yield at least a minimum number of observations per agency and no more than a maximum number. The expected minimum per agency was 15 completed eight-month patient interviews, and the maximum was 100 completed eight-month patient interviews. The sampling fraction was calculated to yield 2,500 completed four-month interviews (and 2,000 completed eight-month interviews) during a sample intake period of six months, after taking into account the agencies for which the minimum and maximum sample sizes applied. For small agencies, the sampling

[2]We considered trying to identify ineligible patients between the four- and eight-month interviews but rejected this approach, on the grounds that the expense of doing so was likely to be greater than the cost of the interviews completed with ineligible patients.

fraction was increased so that the minimum sample size could be achieved. For the smallest agencies, all demonstration patients admitted during the sample intake period were selected. (We dropped one agency before sample selection began because our experience in the first demonstration year indicated that it was unlikely to have more than one or two admissions during the sample intake period.[3]) For the largest agencies, the sampling fraction was reduced so that no more than 100 completed eight-month interviews were expected.

[3]Three agencies had dropped out of the demonstration by the time the patient survey sample was selected.

APPENDIX B

SENSITIVITY ANALYSES

This appendix presents the results of a variety of sensitivity analyses, most of which examine the sensitivity of our results to an alternative weighting scheme. We also examine the effect of removing an outlier agency on the outcome of same-body-system admissions to home health agencies, and the effects of controlling for additional baseline agency characteristics in the regressions for same-body-system hospitalizations.

To test the sensitivity of our results to the weighting scheme, we weighted observations to represent agencies proportional to their size; in contrast, in the main analyses, we weighted observations to represent agencies equally. Analyses in which observations are weighted proportional to size are useful because the "equal weight" analyses may be distorted by a few small agencies that have anomalous results. In the "weighted proportional to size" analysis, more weight is given to larger agencies. In cases in which the size weight analysis yielded substantially different estimates from the main equal weight analysis, we inspected scatter plots of agency-level means versus agency size to identify any small agencies that had atypical results, which could disproportionately affect the equal weight analysis, and to identify any large agencies with atypical results, which could disproportionately affect the size weight analysis.

The following two tables show that, for general health outcomes from the patient survey data (Table B.1) and for improvement and stabilization in basic activities of daily living (ADLs) from the quality assurance (QA) data (Table B.2), neither weighting scheme revealed any real demonstration impacts. Among the numerous comparisons there was only one significant difference in the size weight analysis in satisfaction with life at eight months, favoring treatment agencies (Table B.1). This isolated significant difference probably represents a statistical artifact, and our overall interpretation remains unchanged--there was probably no impact of prospective payment on these outcomes, and certainly no adverse impact.

TABLE B.1

GENERAL HEALTH FROM SURVEY, BY WEIGHTING SCHEME

Outcome	Weighted to Represent Agencies Equally[a]		Weighted to Represent Agencies by Size[b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]
Health Good or Excellent				
Four-month survey	43.0	−0.02	44.7	3.7
		(0.99)		(0.17)
Eight-month survey	46.4	−1.0	45.5	1.0
		(0.73)		(0.75)
Days in Bed in Previous Two Weeks				
Four-month survey	25.2	2.2	24.9	0.9
		(0.34)		(0.71)
Eight-month survey	22.1	−0.8	20.0	−1.3
		(0.76)		(0.63)
Satisfied with Life				
Four-month survey	64.7	2.2	62.3	3.6
		(0.41)		(0.21)
Eight-month survey	62.9	2.9	61.0	4.6**
		(0.21)		(0.03)

SOURCE: Four-month and eight-month patient surveys.

NOTE: The sample includes only first admissions to demonstration home health agencies.

[a]Observations have been weighted to give agencies equal representation in the analysis.

[b]Observations have been weighted to give agencies representation proportional to their size in the analysis.

[c]Regression-adjusted through logit models to control for preexisting differences between treatment and control agencies.

[d]The p-value is based on standard errors inflated to account for the effects of clustering and weighting.

**Significantly different from zero at the 0.05 level, two-tailed test.

TABLE B.2

IMPROVEMENT AND STABILIZATION OUTCOMES IN ACTIVITIES OF DAILY LIVING FROM QUALITY ASSURANCE DATA, BY WEIGHTING SCHEME

Outcome	Weighted to Represent Agencies Equally [a]		Weighted to Represent Agencies by Size [b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference [c] (p-Value) [d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference [c] (p-Value) [d]
Grooming				
Improvement	53.9	-2.2	55.9	-0.5
		(0.33)		(0.80)
Stabilization	87.3	-0.08	88.1	-0.8
		(0.95)		(0.37)
Bathing				
Improvement	51.6	-1.0	55.0	0.4
		(0.67)		(0.84)
Stabilization	84.5	0.47	85.0	-0.6
		(0.68)		(0.55)
Toileting				
Improvement	55.1	-1.3	57.4	-2.4
		(0.60)		(0.24)
Stabilization	91.1	-0.03	91.6	-0.6
		(0.97)		(0.54)
Transferring				
Improvement	46.9	-1.8	50.2	-1.4
		(0.46)		(0.52)
Stabilization	88.1	-1.0	88.9	-1.0
		(0.36)		(0.30)
Ambulating				
Improvement	31.9	-0.9	33.2	0.9
		(0.68)		(0.55)
Stabilization	90.2	-0.48	90.4	-0.4
		(0.61)		(0.62)

SOURCE: Quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies.

[a] Observations have been weighted to give agencies equal representation in the analysis.

[b] Observations have been weighted to give agencies representation proportional to their size in the analysis.

[c] Regression-adjusted through logit models to control for preexisting differences between treatment and control agencies.

[d] The p-value is based on standard errors inflated to account for the effects of clustering and weighting.

However, there were some differences between the equal weight and size weight analyses in the basic ADL outcomes of bathing, eating, transferring, and ambulating obtained from the survey data (Table B.3). In the equal weight analysis, treatment-control differences in did bathe and could bathe at eight months were significant and favored treatment agencies. In the size weight analysis, these differences were smaller and had dropped below the significance level, but their directions and relative magnitudes were roughly comparable.

The apparent negative demonstration effects observed in the equal weight analysis on did eat at four months and eight months, could eat at four months, and did transfer at four months did not persist in the size weight analysis. Scatter plots for the first three outcomes suggest that three small treatment agencies, whose observations were therefore inflated in the equal weight analysis but deflated in the size weight analysis, were responsible for the difference between the two analyses (Figures B.1 and B.2). A few very small treatment agencies were likewise responsible for the discrepancy between size weight and equal weight analyses in did transfer at four months (Figure B.3). The scatter plot for did ambulate at eight months did not contain any small anomalous agencies to explain the discrepancy between the equal weight and size weight analyses (Figure B.3). However, we have no evidence from any other outcome to suggest that we have overlooked a negative impact on functional outcomes in the equal weight analysis.

The key message of Tables B.1 through B.4 and of Figures B.1 through B.3 is that we have no evidence from the survey data of any large or broad-based adverse effects on basic ADL outcomes. Furthermore, the equal weight and size weight analyses of the instrumental ADL outcomes from QA and survey data do not provide any evidence of demonstration impacts (Table B.4).

TABLE B.3

BASIC ACTIVITIES OF DAILY LIVING FROM SURVEY DATA, BY WEIGHTING SCHEME

Outcome	Weighted to Represent Agencies Equally[a]		Weighted to Represent Agencies by Size[b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]
Did Bathe				
Four-month survey	56.5	-1.5	59.2	0.27
		(0.48)		(0.90)
Eight-month survey	56.1	3.4*	61.4	2.9
		(0.09)		(0.16)
Could Bathe				
Four-month survey	65.0	-0.5	67.8	1.2
		(0.81)		(0.57)
Eight-month survey	65.0	3.6**	69.2	2.8
		(0.05)		(0.14)
Did Eat				
Four-month survey	83.5	-3.7*	83.1	-0.7
		(0.06)		(0.64)
Eight-month survey	84.0	-4.1**	83.8	-0.15
		(0.03)		(0.92)
Could Eat				
Four-month survey	92.8	-4.1**	92.7	0.05
		(0.02)		(0.98)
Eight-month survey	91.7	-1.9	91.5	1.7
		(0.16)		(0.20)
Did Transfer				
Four-month survey	70.3	-4.2*	71.0	-2.3
		(0.08)		(0.29)
Eight-month survey	73.9	-1.9	83.8	-0.25
		(0.56)		(0.9)
Could Transfer				
Four-month survey	84.7	-2.0	85.0	-0.77
		(0.26)		(0.62)
Eight-month survey	84.8	-1.0	85.5	0.24
		(0.55)		(0.9)
Did Ambulate				
Four-month survey	70.2	-3.4	71.0	-2.5
		(0.11)		(0.12)
Eight-month survey	85.1	-2.6	87.9	-5.5***
		(0.24)		(0.01)
Could Ambulate				
Four-month survey	82.4	-1.3	82.8	0.07
		(0.57)		(0.98)
Eight-month survey	94.4	-0.23	94.4	-0.73
		(0.87)		(0.57)

SOURCE: Four-month and eight-month patient surveys.

NOTE: The sample includes only first admissions to demonstration home health agencies.

TABLE B.3 *(continued)*

[a] Observations have been weighted to give agencies equal representation in the analysis.

[b] Observations have been weighted to give agencies representation proportional to their size in the analysis.

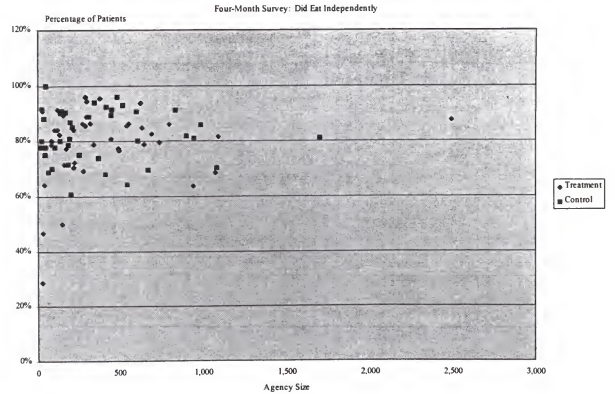
[c] Regression-adjusted through logit models to control for preexisting differences between treatment and control agencies.

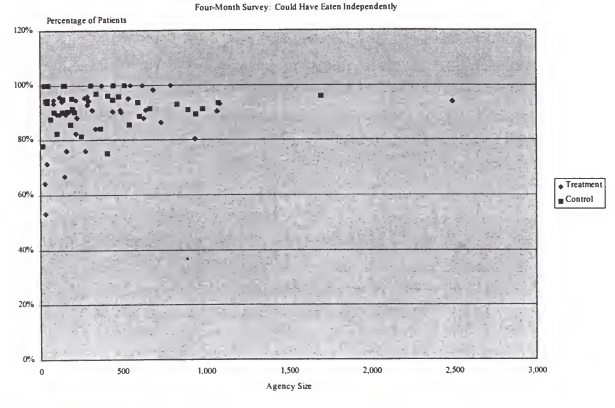
[d] The *p*-value is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

FIGURE B.1

SCATTER PLOTS OF AGENCY-LEVEL MEANS OF SELECTED SURVEY OUTCOMES AGAINST AGENCY SIZE

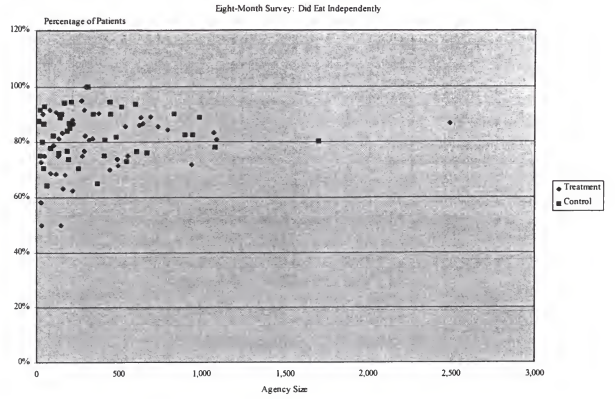




SOURCE: Four-month survey.

FIGURE B.2

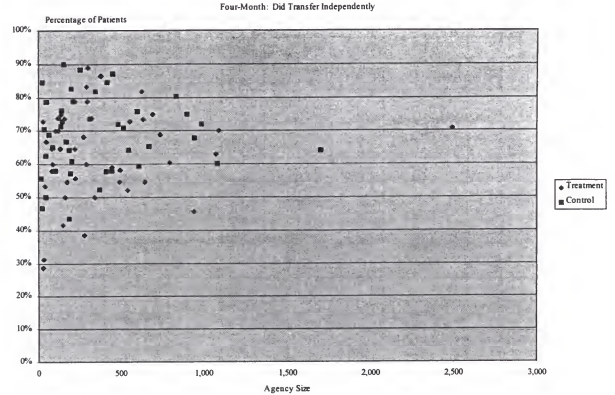
SCATTER PLOT OF AGENCY-LEVEL MEAN OF SELECTED SURVEY OUTCOME AGAINST AGENCY SIZE

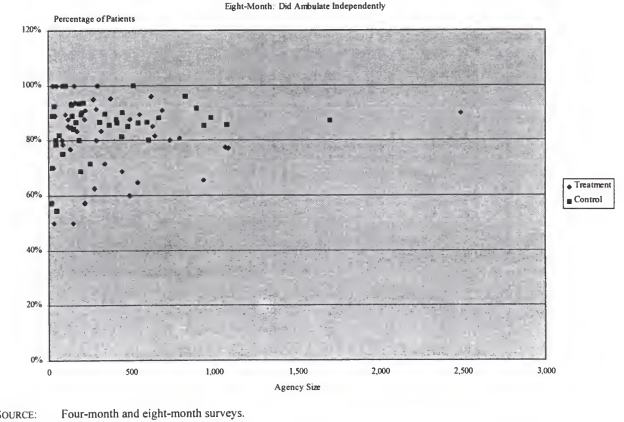


SOURCE: Eight-month survey.

FIGURE B.3

SCATTER PLOTS OF AGENCY-LEVEL MEANS OF SELECTED SURVEY OUTCOMES AGAINST AGENCY SIZE





SOURCE: Four-month and eight-month surveys.

TABLE B.4

INSTRUMENTAL ACTIVITIES OF DAILY LIVING FROM QUALITY ASSURANCE AND SURVEY DATA, BY WEIGHTING SCHEME

Outcome	Weighted to Represent Agencies Equally[a]		Weighted to Represent Agencies by Size[b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]
Quality Assurance Data				
Light Meal Preparation				
Improvement	46.6	-2.5	48.7	-0.3
		(0.27)		(0.88)
Stabilization	89.0	0.6	89.2	0.4
		(0.59)		(0.75)
Housekeeping				
Improvement	41.5	0.2I	42.9	1.3
		(0.93)		(0.50)
Stabilization	75.7	1.8	76.4	0.6
		(0.48)		(0.80)
Management of Medications				
Improvement	36.7	-2.3	38.7	0.00
		(0.37)		(1.00)
Stabilization	87.2	0.7	87.4	0.6
		(0.58)		(0.59)
Survey Data				
Did Take Medications				
Four-month survey	52.4	1.7	55.0	2.6
		(0.44)		(0.17)
Eight-month survey	55.9	1.4	60.7	-0.3
		(0.5)		(0.89)
Could Take Medications				
Four-month survey	68.9	5.0	72.2	3.1
		(0.02)		(0.16)
Eight-month survey	72.6	0.4	73.5	1.0
		(0.9)		(0.57)

SOURCE: Quality assurance data and four-month and eight-month patient surveys.

NOTE: The sample includes only first admissions to demonstration home health agencies.

[a] Observations have been weighted to give agencies equal representation in the analysis.

[b] Observations have been weighted to give agencies representation proportional to their size in the analysis.

[c] Regression-adjusted through logit models to control for preexisting differences between treatment and control agencies.

[d] The p-value is based on standard errors inflated to account for the effects of clustering and weighting.

Table B.5 compares several QA outcomes under both weighting schemes. The negative estimated effect of prospective payment on improvement in surgical wound status is robust to size weighting. In fact, the scatter plot shows three very small *control* group agencies with low rates of surgical wound improvement, which would tend, if anything, to reduce any negative effect of the treatment agencies (Figure B.4). As noted in this report, however, we have no additional corroboration that this isolated negative effect represents a true impact. Significant treatment-control differences favoring treatment agencies in the equal weight analysis, stabilization in urinary incontinence/catheter, and improvement in confusion cease to be significant in the size weight analysis, but again, the directions of the differences are the same, and the magnitudes are roughly comparable. The scatter plots for these two outcomes do not reveal any obvious anomalous agencies (data not shown).

Table B.6 shows that the effects of prospective payment on emergency care and on hospital admission for same-body-system diagnoses were robust to size weighting. Scatter plots for these outcomes did not reveal any obvious outliers that could have distorted the results (data not shown).

In contrast, the results for same-body-system admissions to home health agencies, which were also robust to size weighting, were skewed by a single, large treatment agency that had an unusually high rate of admissions (Figures B.5 and B.6). Both the mean outcome and size of this agency, VNA of Dade County, Florida, were of the right magnitude to distort the results of the equal weight analysis and the size weight analysis.

To assess the effect of this agency on the results, we reanalyzed the data after excluding the agency from the sample. Excluding the agency eliminated the treatment-control difference in both the equal weight and size analysis. As described in Chapter IV, the agency had an exceptionally high rate of loss of its patients to other home health agencies because of an erroneous notice of

TABLE B.5

IMPROVEMENT OUTCOMES FROM QUALITY ASSURANCE DATA, BY WEIGHTING SCHEME

Outcome	Weighted to Represent Agencies Equally[a]		Weighted to Represent Agencies by Size[b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]
Pain				
Improvement	52.2	2.2 (0.28)	55.1	1.5 (0.40)
Stabilization	83.1	1.1 (0.26)	83.3	0.1 (0.87)
Most Problematic Pressure Ulcer				
Improvement	72.6	-1.0 (0.81)	71.8	1.2 (0.75)
Stabilization	97.5	-0.14 (0.69)	97.4	-0.05 (0.84)
Surgical Wound Status				
Improvement	86.2	-3.3** (0.05)	85.1	-4.1** (0.04)
Stabilization	97.4	0.17 (0.83)	97.7	0.06 (1.00)
Dyspnea				
Improvement	45.3	1.4 (0.57)	49.6	2.6 (0.13)
Stabilization	79.2	1.8 (0.18)	79.7	1.0 (0.39)
Urinary Tract Infection				
Improvement	80.9	-2.6 (0.44)	82.1	-3.1 (0.27)
Stabilization	98.1	0.38 (0.22)	98.2	0.02 (0.95)
Urinary Incontinence or Catheter Present				
Improvement	43.9	2.5 (0.38)	49.0	4.6* (0.06)
Stabilization	93.8	1.1** (0.03)	95.0	0.6 (0.14)
Confusion				
Improvement	37.7	5.7* (0.05)	42.5	4.7 (0.11)
Stabilization	84.9	2.0 (0.12)	86.5	0.9 (0.32)
Behavior Problem Frequency				
Improvement	62.5	-1.2 (0.72)	64.1	2.9 (0.26)
Stabilization	91.5	0.72 (0.42)	92.3	0.7 (0.35)

TABLE B.5 *(continued)*

Outcome	Weighted to Represent Agencies Equally[a]		Weighted to Represent Agencies by Size[b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (*p*-Value)[d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (*p*-Value)[d]
Mortality				
Within 120 Days	9.5	0.1 (2.85)	10.0	−0.07 (0.17)
Within Eight Months	15.4	−0.2 (0.8)	16.0	−1.0 (0.14)
Within One Year	20.3	−0.7 (0.32)	20.9	−1.5** (0.03)

SOURCE: Quality assurance data.

NOTE: The sample includes only first admissions to demonstration home health agencies.

[a]Observations have been weighted to give agencies equal representation in the analysis.

[b]Observations have been weighted to give agencies representation proportional to their size in the analysis.

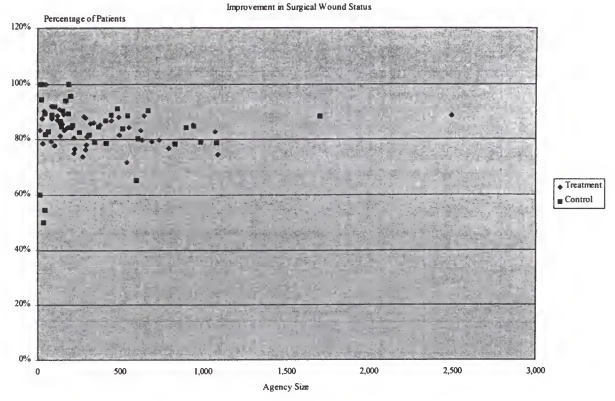
[c]Regression-adjusted through logit models to control for preexisting differences between treatment and control agencies.

[d]The *p*-value is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.

FIGURE B.4

SCATTER PLOT OF AGENCY-LEVEL MEAN OF SELECTED QUALITY ASSURANCE OUTCOMES AGAINST AGENCY SIZE



SOURCE: Quality assurance data.

TABLE B.6

HEALTH SERVICES USE FROM QUALITY ASSURANCE AND CLAIMS DATA, BY WEIGHTING SCHEME

Outcome	Weighted to Represent Agencies Equally[a]		Weighted to Represent Agencies by Size[b]	
	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference[c] (p-Value)[d]
Quality Assurance Data				
Emergency Care				
Hospital emergency room	10.7	−0.9	9.3	−1.5*
		(0.3)		(0.09)
Outpatient clinic or urgent care clinic	1.3	−0.5***	1.0	−0.34**
		(0.01)		(0.05)
Physician's office	3.0	−1.0***	2.3	−1.0***
		(0.01)		(0.004)
Any of the above	13.9	−2.0*	11.9	−2.2**
		(0.08)		(0.03)
Medicare Claims Data				
Admission to Hospital for Same-Body-System Diagnosis				
Within 120 days	21.2	−1.1	20.6	−0.78
		(0.20)		(0.16)
Within eight months	29.4	−2.0*	28.7	−1.6**
		(0.06)		(0.03)
Within one year	35.2	−2.4**	34.3	−1.5*
		(0.04)		(0.07)
Admission to SNF for Same-Body-System Diagnosis				
Within 120 days	5.8	−0.5	5.8	−0.5*
		(0.22)		(0.06)
Within eight months	9.1	−0.5	8.8	−0.6*
		(0.38)		(0.09)
Within one year	11.5	−0.4	11.3	−0.7
		(0.49)		(0.14)
Admission to HHA for Same-Body-System Diagnosis				
Within 120 days	5.6	1.1*	5.0	1.1**
		(0.09)		(0.04)
Within eight months	11.1	1.8*	10.6	1.8**
		(0.06)		(0.01)
Within one year	17.2	2.4**	17.1	1.6*
		(0.03)		(0.08)

SOURCE: Quality assurance and Medicare claims data.

NOTE: The sample includes only first admissions to demonstration home health agencies.

[a] Observations have been weighted to give agencies equal representation in the analysis.

[b] Observations have been weighted to give agencies representation proportional to their size in the analysis.

TABLE B.6 *(continued)*

[c] Regression-adjusted through logit models to control for preexisting differences between treatment and control agencies.

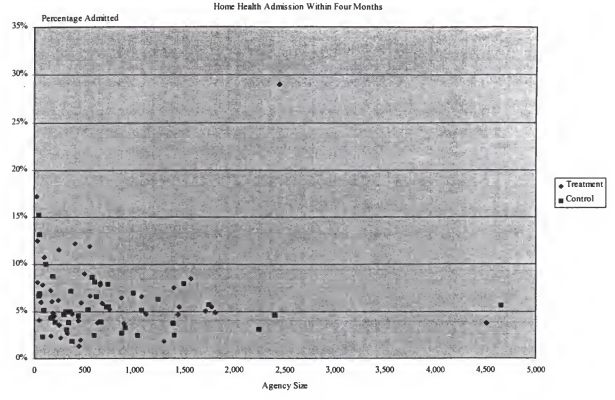
[d] The *p*-value is based on standard errors inflated to account for the effects of clustering and weighting.

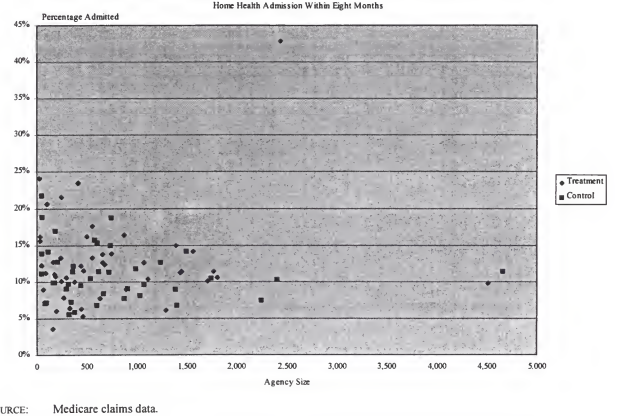
SNF = skilled nursing facility; HHA = home health agency.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

FIGURE B.5

SCATTER PLOTS OF AGENCY-LEVEL MEAN RATES OF HOME HEALTH ADMISSION FOR SAME-BODY-SYSTEM DIAGNOSIS AT FOUR AND EIGHT MONTHS AGAINST AGENCY SIZE

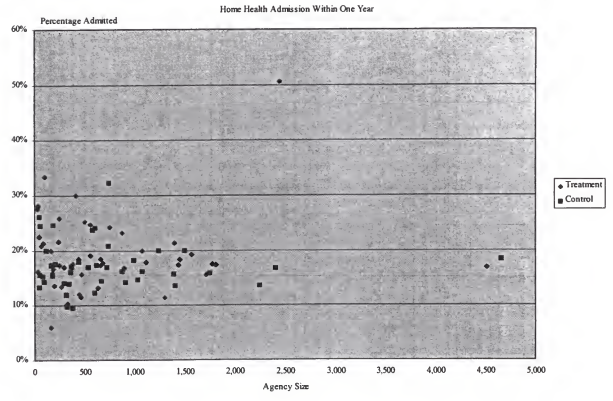




SOURCE: Medicare claims data.

FIGURE B.6

SCATTER PLOT OF AGENCY-LEVEL MEAN RATE OF HOME HEALTH ADMISSION FOR SAME-BODY-SYSTEM DIAGNOSIS AT ONE YEAR AGAINST AGENCY SIZE



SOURCE: Medicare claims data.

closure. Thus, we conclude that prospective payment did not have an impact on same-body-system admissions to home health agencies.

The equal weight and size weight analyses for the satisfaction outcomes from the patient survey did not yield any unexpected findings. Neither weighting scheme showed any impact on the outcomes of general or overall satisfaction with agency care (data not shown). The demonstration effects on the three satisfaction outcomes in the equal weight analysis (staff rushed through work most or all of the time, disagree or strongly disagree that staff encouraged independence, and staff paid attention to the patient some of the time or little or none of the time) persisted in the size weight analysis (data not shown). Scatterplots for these three outcomes did not identify any outlier agencies responsible for the observed effects (data not shown).

As discussed in Chapters IV and V, we examined the possibility that the apparent effects of prospective payment on same-body-system hospitalizations in the main effects analysis and the for-profit subgroup analysis could have resulted from an underlying difference between treatment and control agencies in baseline hospitalization rates. First, we performed a regression that included a new control variable: the agency-level percentage of episodes with a hospitalization during the base-quarter. Second, we excluded the four agencies with the highest value for this variable (all four were control group, for-profit agencies from Illinois). Compared with the original analyses, both additional analyses produced smaller and nonsignificant treatment-control differences but did not eliminate them entirely (Tables B.7 and B.8). There is thus evidence that uncontrolled differences between treatment and control agencies may partly account for the reductions in same-body-system hospitalizations found in the original main effects and for-profit subgroup analyses.

TABLE B.7

THREE DIFFERENT REGRESSIONS TO ESTIMATE MAIN TREATMENT-CONTROL EFFECTS ON SAME-BODY-SYSTEM DIAGNOSIS HOSPITALIZATIONS

Hospital Admission for Same-Body-System Diagnosis	Unadjusted Control Group Mean (Percentage)	Estimated Treatment-Control Difference (p-Value)[a]
"Standard" Regression[b] (N = 65,284)		
At 120 Days	21.2	-1.1 (0.20)
At Eight Months	29.4	-2.0** (0.06)
At One Year	35.2	-2.4** (0.04)
With the New Control Variable: Agency-Based Percentage of Episodes with a Hospitalization in Base Quarter[c] (N = 65,284)		
At 120 Days	21.2	-0.73 (0.39)
At Eight Months	29.4	-1.4 (0.17)
At One Year	35.2	-1.6 (0.14)
After Removing Four Illinois For-Profit Control Agencies[d] (N = 63,097)		
At 120 Days	20.3	-0.61 (0.46)
At Eight Months	28.6	-1.4 (0.18)
At One Year	34.2	-1.6 (0.14)

SOURCE: Medicare claims data.

NOTE: The sample includes only first admissions to demonstrate home health agencies. Observations are weighted to give agencies equal representation.

[a] The p-value corresponds to a test of whether the treatment-control difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[b] Includes all the standard control variables listed in Table II.5.

[c] Includes standard control variables plus the new control variable of agency-based percentage of episodes with a hospitalization during the base quarter.

[d] After removing the four for-profit control agencies from Illinois with unusually high percentages of episodes with a hospitalization in the base quarter.

** Significantly different from zero at the 0.05 level, two-tailed test.

TABLE B.8

THREE DIFFERENT REGRESSIONS TO ESTIMATE SUBGROUP TREATMENT-CONTROL EFFECTS ON SAME-BODY-SYSTEM DIAGNOSIS HOSPITALIZATIONS, FOR-PROFIT OR NONPROFIT SUBGROUP

Hospital Admission for Same-Body-System Diagnosis	For Profit			Nonprofit			*p*-Value for Difference in Subgroup Effect[b]
	Control Group Mean (Percentage)	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	Control Group Mean (Percentage)	Estimated Treatment-Control Difference (*p*-Value[a])	Estimated Difference as Percentage of Control Group Mean	
"Standard" Regression[c] (N = 65,284)							
At 120 Days	22.2	−3.9*** (0.00)	−17.6	20.1	1.6 (0.25)	8.0	0.01***
At Eight Months	30.4	−4.8*** (0.00)	−15.8	28.4	0.4 (0.79)	1.4	0.04**
At One Year	36.2	−4.8*** (0.01)	−13.3	34.2	−0.1 (0.94)	−0.3	0.06*
With the New Control Variable: Agency-Based Percentage of Episodes with a Hospitalization in Base Quarter[d] (N = 65,284)							
At 120 Days	22.0	−4.0*** (0.01)	−18.2	19.0	2.0 (0.16)	10.5	0.01***
At Eight Months	30.6	−4.2*** (0.01)	−13.7	27.5	1.1 (0.51)	4.0	0.04**
At One Year	31.9	−4.1** (0.02)	−12.9	33.2	0.7 (0.64)	2.1	0.07*
After Removing Four Illinois For-Profit Control Agencies[e] (N = 63,097)							
At 120 Days	18.4	−3.0 (0.03)	−16.3	19.6	1.3 (0.37)	6.6	0.06*
At Eight Months	29.7	−3.6 (0.03)	−12.1	28.2	0.0 (1.0)	0.0	0.17
At One Year	34.9	−3.5 (0.06)	−10.0	34.1	−0.7 (0.71)	−2.1	0.27

TABLE B.8 *(continued)*

SOURCE: Medicare claims data.

NOTE: The sample includes only first admissions to demonstratc home health agencies. Observations have been weighted to give agencies equal representation. *p*-Values are based on standard errors that account for the effects of clustering and weighting.

[a] This *p*-value is the significance level for the hypothesis of a demonstration effect *within* the subgroup.

[b] This *p*-value is the significance level for the hypothesis of a subgroup effect (that is, a difference in impacts across subgroups). It is the *p*-value for the treatment status by subgroup variable interaction term in the regression.

[c] Includes all the standard control variables listed in Table II.5.

[d] Includes standard control variables plus the new control variable of agency-based percentage of episodes with a hospitalization during the base quarter.

[e] After removing the four for-profit control agencies from Illinois with unusually high percentages of episodes with a hospitalization during the base quarter.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

APPENDIX C

COEFFICIENT ESTIMATES FROM SELECTED REPRESENTATIVE REGRESSIONS

TABLE C.1

ESTIMATED COEFFICIENTS FROM LOGISTIC REGRESSIONS MODELS OF ADMISSION TO HOSPITAL FOR SAME BODY SYSTEM DIAGNOSIS

	Admission to Hospital for Same-Body-System Diagnosis		
	Within 120 Days (*p*-Value)	Within Eight Months (*p*-Value)	Within One Year (*p*-Value)
Treatment/Control Status			
Agency Was Prospectively Paid	-0.07 (0.20)	-0.11* (0.06)	-0.11** (0.04)
Demographic Measures			
Original Reason for Medicare: Reached Age 65 Years	-0.14** (0.04)	-0.18*** (0.01)	-0.19*** (0.01)
Age (Years)			
Younger than 65	-0.10 (0.31)	-0.06 (0.53)	0.06 (0.56)
75 to 84	0.07 (0.15)	0.10** (0.04)	0.13*** (0.01)
85 or older	-0.03 (0.52)	0.04 (0.41)	0.09* (0.08)
White	0.11* (0.10)	0.11* (0.09)	0.12* (0.09)
Female	-0.12*** (0.00)	-1.0** (0.02)	-0.08** (0.04)
Has Medicaid Buy-In for Part A and B Medicare	0.11** (0.03)	0.16*** (0.00)	0.15*** (0.00)
Enrolled in Medicare for Less Than Six Months Before Home Health Admission	0.49*** (0.01)	0.34** (0.02)	0.22* (0.10)
Enrolled in an HMO at Some Time Six Months Before Home Health Admission	0.08 (0.65)	0.26 (0.26)	0.29* (0.10)
Had Medicare as Secondary Payer at Some Time Six Months Before Home Health Admission	-0.56** (0.05)	-0.50** (0.02)	-0.66*** (0.00)
Medical Conditions, Symptoms, and Needs at Home Health Admission			
Cancer	0.35*** (0.00)	0.28*** (0.00)	0.19*** (0.00)
Diabetes	0.29*** (0.00)	0.32*** (0.00)	0.33*** (0.00)
Cerebrovascular Accident	-0.02 (0.67)	0.04 (0.29)	0.06* (0.09)
Stage 3 or 4 Decubitus Ulcers	0.11 (0.21)	0.13* (0.08)	0.09 (0.21)

TABLE C.1 *(continued)*

	Admission to Hospital for Same-Body-System Diagnosis		
	Within 120 Days (p-Value)	Within Eight Months (p-Value)	Within One Year (p-Value)
Need for Complicated Wound Care	-0.06 (0.43)	-0.10* (0.07)	-0.06 (0.18)
Functional Limitations			
Bathing	0.02 (0.76)	-0.00[a] (0.95)	0.05 (0.29)
Eating	0.12** (0.03)	0.09* (0.07)	0.07 (0.12)
Dressing	0.01 (0.92)	0.04 (0.38)	-0.01 (0.78)
Toileting	0.07 (0.19)	0.10 (0.85)	-0.01 (0.87)
Transferring	-0.07 (0.12)	-0.06 (0.19)	-0.05 (0.19)
Measures of Patient Prior Service Use Before Home Health Admission			
Was in Hospital Before Home Health Admission	-0.01 (0.93)	-0.07 (0.12)	-0.07 (0.14)
Length of Inpatient Stay Ending in Two Weeks Before Home Health Admission (Days)	0.00[a] (0.16)	0.00[a] (0.22)	0.01** (0.02)
Whether in SNF in Two Weeks Before Home Health Admission	-0.08 (0.27)	-0.04 (0.60)	-0.00[a] (0.96)
Home Health Visits from Nondemonstration Agencies in Six Months Before Home Health Admission (Number)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Hospitalizations in Six Months Before Home Health Admission (Number)	0.34*** (0.00)	0.35*** (0.00)	0.35*** (0.00)
SNF Admissions in Six Months Before Home Health Admission (Number)	-0.10* (0.09)	-0.10** (0.05)	-0.10** (0.03)
Time of Home Health Admission			
Quarter Admitted to Home Health Care			
Third quarter of calendar year 1995	-0.03 (0.72)	-0.03 (0.71)	-0.06 (0.43)
Fourth quarter of calendar year 1995	-0.05 (0.35)	-0.01 (0.89)	0.02 (0.78)
Second quarter 1996	-0.14*** (0.01)	-0.05 (0.24)	-0.04 (0.39)
Third quarter 1996	-0.06 (0.24)	-0.01 (0.92)	-0.02 (0.73)
Fourth quarter 1996	-0.09 (0.17)	-0.03 (0.58)	-0.06 (0.31)

TABLE C.1 *(continued)*

	Admission to Hospital for Same-Body-System Diagnosis		
	Within 120 Days (*p*-Value)	Within Eight Months (*p*-Value)	Within One Year (*p*-Value)
Admitted in Agency's Second or Third Demonstration Year	0.13 (0.19)	0.13 (0.16)	0.11 (0.16)
Interaction Term of Treatment Status Times Admitted in Agency's Second or Third Demonstration Year	0.04 (0.77)	0.02 (0.87)	0.01 (0.89)
Agency Characteristics			
For Profit	0.03 (0.74)	0.02 (0.85)	0.01 (0.94)
Hospital Based	0.00* (0.98)	−0.02 (0.80)	−0.01 (0.87)
Member of a Chain	0.02 (0.75)	0.06 (0.37)	0.07 (0.23)
Provided Fewer than 30,000 Visits in Base Year	0.05 (0.56)	0.04 (0.62)	0.01 (0.85)
Predemonstration Practice Pattern (Ratio)[b]	−0.05 (0.67)	−0.10 (0.43)	−0.05 (0.66)
Area Characteristics			
State			
Florida	0.04 (0.47)	0.08 (0.14)	0.11* (0.07)
Illinois	0.14 (0.14)	0.23*** (0.01)	0.22*** (0.01)
Massachusetts	0.17 (0.16)	0.28** (0.02)	0.25** (0.04)
Texas	−0.02 (0.81)	0.05 (0.54)	0.06 (0.49)
Urban Area	0.16 (0.22)	0.04 (0.73)	0.09 (0.36)
Physicians per 10,000 Residents, 1994 (Number)	−0.00* (0.35)	−0.00* (0.61)	−0.00* (0.69)
Nursing Home Beds per 100 Residents Older than Age 65, 1991 (Number)	0.02 (0.32)	0.01 (0.80)	0.01 (0.51)
Hospital Occupancy Rate, 1993	0.14 (0.74)	0.15 (0.66)	0.18 (0.56)

TABLE C.1 *(continued)*

	Admission to Hospital for Same-Body-System Diagnosis		
	Within 120 Days (p-Value)	Within Eight Months (p-Value)	Within One Year (p-Value)
Mean Medicare Reimbursement per Beneficiary, 1991 (Dollars)	-0.04 (0.57)	-0.07 (0.33)	-0.06 (0.34)
Intercept	-1.85*** (0.00)	-1.24*** (0.00)	-1.15*** (0.00)

SOURCE: Medicare claims data.

NOTE: Observations have been weighted to represent agencies equally, and p-values have been adjusted for the effects of weighting and clustering. Patients with a previous demonstration home health admission within the year preceding the index admission have been excluded.

[a]Coefficient has absolute value of less than 0.005.

[b]An index of the case-mix-adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

SNF = skilled nursing facility.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

TABLE C.2

ESTIMATED COEFFICIENTS FROM LOGISTIC REGRESSION MODELS OF FUNCTIONAL STATUS OUTCOMES IN QUALITY ASSURANCE DATA

	Grooming		Bathing		Toileting	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Treatment/Control Status						
Agency Was Prospectively Paid	-0.10 (0.33)	-0.01 (0.94)	-0.04 (0.68)	0.03 (0.71)	-0.06 (0.60)	-0.01 (0.95)
Demographic Measures						
Original Reason for Medicare: Reached Age 65 Years	0.14* (0.09)	-0.06 (0.47)	0.22*** (0.00)	0.12 (0.24)	-0.03 (0.78)	0.12 (0.21)
Age (Years)						
Younger than 65	-0.01 (0.92)	-0.10 (0.44)	0.00* (0.95)	0.06 (0.64)	-0.27* (0.07)	-0.03 (0.88)
75 to 84	-0.14*** (0.01)	-0.12** (0.02)	-0.20** (0.00)	-0.09* (0.08)	0.04 (0.53)	-0.09 (0.16)
85 or older	-0.38*** (0.00)	-0.42*** (0.00)	-0.46*** (0.00)	-0.37*** (0.00)	-0.20** (0.02)	-0.38*** (0.00)
White	0.06 (0.36)	-0.10 (0.20)	-0.05 (0.50)	-0.05 (0.45)	-0.06 (0.45)	-0.04 (0.62)
Female	0.13*** (0.01)	0.04 (0.33)	0.04 (0.26)	0.04 (0.31)	0.03 (0.53)	0.02 (0.77)
Has Medicaid Buy-In for Part A and B Medicare	-0.20*** (0.00)	-0.05 (0.40)	-0.17*** (0.01)	0.00* (0.95)	-0.12 (0.12)	-0.06 (0.38)
Medical Conditions, Symptoms, and Needs at Home Health Admission						
Cancer	-0.18*** (0.01)	-0.39*** (0.00)	-0.20*** (0.00)	-0.35*** (0.00)	-0.17* (0.06)	0.52*** (0.00)
Diabetes	-0.04 (0.39)	-0.08 (0.17)	-0.13*** (0.00)	-0.06 (0.30)	-0.05 (0.59)	0.07 (0.34)
Cerebrovascular Accident	-0.26*** (0.00)	-0.08 (0.19)	-0.22*** (0.00)	-0.14*** (0.01)	0.29*** (0.00)	0.15* (0.07)

TABLE C.2 *(continued)*

	Grooming		Bathing		Toileting	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Need for Complicated Wound Care	0.03 (0.72)	-0.16** (0.04)	0.07 (0.27)	-0.14* (0.07)	-0.17 (0.13)	-0.02 (0.81)
Functional Limitations						
Eating	-0.39*** (0.00)	-0.17*** (0.01)	-0.20*** (0.00)	-0.25*** (0.00)	-0.49*** (0.00)	-0.37*** (0.00)
Dressing	-0.41*** (0.00)	-0.19*** (0.00)	-0.12*** (0.00)	0.20*** (0.00)	-0.39*** (0.00)	-0.70*** (0.00)
Has Risk Factors[b]	0.10* (0.09)	0.09* (0.10)	0.02 (0.70)	0.14*** (0.01)	0.28*** (0.00)	0.18** (0.02)
Unknown Risk Factors[c]	-0.05 (0.54)	0.01 (0.85)	0.01 (0.86)	0.01 (0.82)	0.00* (0.96)	-0.07 (0.37)
Medically Unstable	-0.02 (0.65)	-0.04 (0.48)	0.02 (0.54)	-0.08* (0.06)	0.02 (0.71)	-0.17*** (0.01)
Depressed Feelings	0.15*** (0.01)	0.17** (0.02)	0.03 (0.51)	0.09 (0.14)	0.24*** (0.00)	0.01 (0.92)
Displays Depressive Behaviors	-0.10 (0.17)	-0.14** (0.05)	0.02 (0.64)	-0.01 (0.85)	-0.16** (0.04)	0.07 (0.35)
Demonstrates Disruptive Behaviors	-0.34*** (0.00)	-0.13* (0.06)	-0.32*** (0.00)	-0.13* (0.09)	-0.37*** (0.00)	-0.16* (0.08)
Prognosis at Home Health Admission						
Likelihood that Treatment Can Be Taken Over[d]	-0.01 (0.81)	0.03 (0.52)	0.02 (0.68)	-0.01 (0.79)	0.00* (0.99)	-0.08 (0.14)
Prognosis Is Good/Fair	0.33*** (0.00)	0.31*** (0.00)	0.33*** (0.00)	0.33*** (0.00)	0.44*** (0.00)	0.43*** (0.00)
Life Expectancy Less than Six Months	-0.41*** (0.00)	-0.26*** (0.00)	-0.16** (0.04)	-0.32*** (0.00)	-0.33*** (0.00)	-0.49*** (0.00)
Rehabilitative Prognosis Is Good	0.50*** (0.00)	0.51*** (0.00)	0.41*** (0.00)	0.40*** (0.00)	0.48*** (0.00)	0.41*** (0.00)

TABLE C.2 *(continued)*

	Grooming		Bathing		Toileting	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Availability of Informal Care at Home Health Admission						
Live-In Informal Help	-0.31***	-0.24***	-0.16***	-0.12**	-0.25***	-0.29***
	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)
Paid Help or in Assisted-Living Residence	-0.52***	-0.35***	-0.42***	-0.26***	-0.51***	-0.32***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Measures of Patient's Service Use Before Home Health Admission						
Was in Hospital Before Home Health Admission	0.25***	0.28***	0.28***	0.23***	0.21***	0.20***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Length of Inpatient Stay Ending in Two Weeks Before Home Health Admission (Days)	0.01***	0.00*	0.01***	0.01	0.01	0.00*
	(0.00)	(0.61)	(0.00)	(0.15)	(0.18)	(0.73)
Whether in SNF in Two Weeks Before Home Health Admission	0.19***	0.06	0.14***	0.07	0.18***	0.06
	(0.00)	(0.30)	(0.00)	(0.23)	(0.01)	(0.41)
Home Health Visits from Nondemonstration Agencies in Six Months Before Home Health Admission (Number)	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Hospitalizations in Six Months Before Home Health Admission (Number)	-0.05**	0.03	0.03*	0.04	0.03	0.03
	(0.03)	(0.30)	(0.09)	(0.11)	(0.28)	(0.34)
Time of Home Health Admission						
Quarter Admitted to Home Health Care						
Third quarter 1996	-0.09	-0.08	-0.12	-0.01	-0.21**	0.05
	(0.27)	(0.36)	(0.11)	(0.88)	(0.03)	(0.60)
Fourth quarter 1996	-0.12*	-0.08	-0.11	-0.01	-0.19*	0.03
	(0.09)	(0.44)	(0.16)	(0.87)	(0.08)	(0.81)
First quarter 1997	0.02	0.13	-0.09	-0.17	-0.23	0.25*
	(0.89)	(0.40)	(0.41)	(0.11)	(0.11)	(0.10)
Second quarter 1997	-0.02	0.19	-0.14	0.21*	-0.38***	0.22
	(0.88)	(0.24)	(0.20)	(0.07)	(0.00)	(0.13)
Third quarter 1997	-0.02	0.13	-0.21**	0.11	-0.40***	0.17
	(0.87)	(0.38)	(0.05)	(0.26)	(0.00)	(0.33)

TABLE C.2 *(continued)*

	Grooming		Bathing		Toileting	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Admitted in Agency's Second or Third Demonstration Year	-0.03 (0.76)	-0.15 (0.27)	-0.01 (0.92)	-0.17* (0.08)	0.14 (0.28)	-0.17 (0.16)
Interaction Term of Treatment Status Times Admitted in Agency's Second or Third Demonstration Year	0.09 (0.36)	-0.07 (0.54)	0.14 (0.20)	-0.01 (0.90)	0.03 (0.83)	-0.11 (0.36)
Agency Characteristics						
For Profit	0.07 (0.59)	0.09 (0.49)	-0.04 (0.72)	0.15 (0.18)	-0.06 (0.64)	0.09 (0.55)
Hospital Based	-0.20 (0.14)	0.17* (0.09)	-0.23** (0.04)	0.21** (0.05)	-0.21** (0.15)	-0.00[a] (1.00)
Member of a Chain	-0.18** (0.04)	-0.17 (0.11)	-0.09 (0.29)	-0.11 (0.24)	-0.22** (0.03)	-0.19 (0.13)
Provided Fewer than 30,000 Visits in Base Year	-0.07 (0.59)	-0.03 (0.82)	-0.14 (0.25)	-0.11 (0.33)	-0.02 (0.89)	-0.05 (0.72)
Predemonstration Practice Pattern (Ratio)[f]	-0.23 (0.22)	-0.05 (0.77)	-0.01 (0.97)	0.04 (0.85)	0.02 (0.92)	-0.17 (0.52)
Area Characteristics						
State						
Florida	-0.06 (0.66)	0.28*** (0.01)	-0.08 (0.51)	0.35*** (0.00)	-0.06 (0.66)	0.21 (0.12)
Illinois	0.17 (0.22)	0.18 (0.29)	0.12 (0.44)	0.10 (0.45)	0.06 (0.70)	0.09 (0.62)
Massachusetts	0.34** (0.03)	0.26* (0.08)	0.17 (0.26)	0.11 (0.47)	0.24 (0.13)	0.26 (0.18)
Texas	-0.04 (0.76)	-0.03 (0.81)	-0.02 (0.91)	-0.01 (0.97)	0.07 (0.62)	0.14 (0.35)
Physicians per 10,000 Residents, 1994 (Number)	-0.01* (0.07)	-0.01* (0.09)	-0.01 (0.22)	-0.00[a] (0.44)	-0.01 (0.12)	-0.02** (0.02)
Nursing Home Beds per 100 Residents Older than Age 65, 1991 (Number)	0.00[a] (0.90)	-0.01 (0.80)	-0.02 (0.41)	-0.01 (0.79)	0.01 (0.74)	0.02 (0.44)

TABLE C.2 *(continued)*

	Grooming		Bathing		Toileting	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Hospital Occupancy Rate, 1993	0.14 (0.75)	-0.11 (0.83)	0.48 (0.43)	-0.45 (0.38)	0.65 (0.28)	0.50 (0.40)
Mean Medicare Reimbursement per Beneficiary, 1991 (Dollars)	-0.00[a] (0.81)	0.00[a] (0.37)	-0.00[a] (0.41)	0.00[a] (0.58)	-0.00[a] (0.13)	0.00** (0.06)
Intercept	0.94* (0.09)	1.88*** (0.00)	0.31 (0.54)	1.29*** (0.01)	1.28** (0.03)	2.04*** (0.00)

SOURCE: Quality assurance data.

NOTE: Observations have been weighted to represent agencies equally, and *p*-values have been adjusted for the effects of weighting and clustering. Patients with a previous demonstration home health admission within the year preceding the index admission have been excluded.

[a] Coefficient has absolute value of less than 0.005.

[b] Has any of the following risk factors: heavy smoking, obesity, alcoholism, or drug dependency.

[c] A zero or one indicator variable with the value one if risk factor information was missing. If risk factor information was missing, an imputed value was used for "has risk factors."

[d] By the patient, or by relatives, friend, neighbors, or paid helpers of the patient.

[e] An index of the case-mix-adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

SNF = skilled nursing facility.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

TABLE C.3

ESTIMATED COEFFICIENTS FROM LOGISTIC REGRESSION MODELS OF SELECTED OUTCOMES IN QUALITY ASSURANCE DATA

	Transferring		Oral Medication Management		Pain	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Treatment/Control Status						
Agency Was Prospectively Paid	-0.08 (0.45)	-0.10 (0.35)	-0.11 (0.37)	0.07 (0.58)	0.09 (0.29)	0.08 (0.27)
Demographic Measures						
Original Reason for Medicare: Reached Age 65 Years	0.19*** (0.01)	0.15* (0.06)	0.08 (0.31)	0.15* (0.07)	0.15*** (0.01)	0.03 (0.63)
Age (Years)						
Younger than 65	0.10 (0.23)	0.06 (0.55)	0.05 (0.65)	0.07 (0.56)	0.08 (0.35)	0.09 (0.35)
75 to 84	-0.15*** (0.00)	-0.07* (0.09)	-0.23*** (0.00)	-0.23*** (0.00)	-0.07 (0.11)	-0.01 (0.90)
85 or older	-0.38*** (0.00)	-0.24*** (0.00)	-0.49*** (0.00)	-0.63*** (0.00)	-0.07 (0.20)	-0.05 (0.38)
White	-0.03 (0.59)	0.00[c] (0.98)	0.03 (0.63)	0.00[c] (0.99)	0.03 (0.63)	-0.09 (0.15)
Female	-0.05 (0.18)	0.01 (0.78)	0.11** (0.02)	0.18*** (0.00)	-0.06* (0.07)	-0.01 (0.88)
Has Medicaid Buy-In for Part A and B Medicare	-0.03 (0.58)	0.00[c] (1.00)	-0.13** (0.04)	-0.03 (0.68)	-0.05 (0.28)	0.08 (0.13)
Medical Conditions, Symptoms, and Needs at Home Health Admission						
Cancer	-0.21*** (0.00)	-0.47*** (0.00)	-0.01 (0.90)	-0.45*** (0.00)	-0.19*** (0.00)	0.26*** (0.00)
Diabetes	-0.07 (0.13)	-0.11** (0.03)	-0.15*** (0.00)	-0.08* (0.09)	-0.11*** (0.01)	-0.08 (0.15)
Cerebrovascular Accident	-0.15*** (0.01)	-0.10* (0.07)	-0.24*** (0.00)	-0.02 (0.84)	0.00[c] (0.99)	0.12** (0.03)

TABLE C.3 *(continued)*

	Transferring		Oral Medication Management		Pain	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Need for Complicated Wound Care	−0.02 (0.80)	−0.04 (0.62)	0.04 (0.63)	0.05 (0.64)	0.13** (0.05)	−0.09* (0.09)
Functional Limitations						
Eating	−0.09** (0.04)	−0.20*** (0.00)	−0.26*** (0.00)	−0.21*** (0.00)	−0.07 (0.14)	−0.20*** (0.00)
Dressing	−0.18*** (0.00)	−0.13** (0.04)	−0.17*** (0.00)	−0.20*** (0.00)	−0.03 (0.36)	−0.03 (0.52)
Has Risk Factors[a]	−0.03 (0.66)	0.16** (0.03)	0.15*** (0.01)	0.13 (0.11)	−0.03 (0.46)	0.10** (0.04)
Unknown Risk Factors[b]	−0.08 (0.26)	0.02 (0.82)	0.08 (0.29)	−0.05 (0.62)	−0.02 (0.78)	0.14* (0.10)
Medically Unstable	0.02 (0.63)	−0.15*** (0.00)	0.01 (0.81)	−0.26*** (0.00)	−0.06 (0.18)	0.01 (0.77)
Depressed Feelings	0.08 (0.19)	0.09 (0.11)	0.09* (0.10)	0.01 (0.93)	−0.01 (0.79)	−0.07 (0.16)
Displays Depressive Behaviors	−0.01 (0.88)	−0.01 (0.90)	−0.11 (0.14)	0.00[c] (0.99)	0.00[c] (0.99)	0.06 (0.34)
Demonstrates Disruptive Behaviors	−0.16*** (0.00)	−0.05 (0.48)	−0.56*** (0.00)	−0.33*** (0.00)	0.04 (0.53)	0.15** (0.02)
Prognosis at Home Health Admission						
Likelihood that Treatment Can Be Taken Over[b]	0.04 (0.46)	−0.01 (0.85)	0.07 (0.20)	−0.04 (0.56)	0.07 (0.23)	0.09* (0.06)
Prognosis Is Good/Fair	0.28*** (0.00)	0.24** (0.02)	0.37*** (0.00)	0.31** (0.02)	0.11 (0.11)	−0.01 (0.89)
Life Expectancy Less than Six Months	−0.19** (0.03)	−0.34*** (0.00)	−0.08 (0.34)	−0.34*** (0.00)	−0.16*** (0.01)	0.04 (0.55)
Rehabilitative Prognosis Is Good	0.42*** (0.00)	0.41*** (0.00)	0.29*** (0.00)	0.21*** (0.00)	0.12** (0.04)	0.20*** (0.01)

TABLE C.3 *(continued)*

	Transferring		Oral Medication Management		Pain	
	Improvement (p-Value)	Stabilization (p-Value)	Improvement (p-Value)	Stabilization (p-Value)	Improvement (p-Value)	Stabilization (p-Value)
Availability of Informal Care at Home Health Admission						
Live-In Informal Help	−0.16*** (0.00)	0.00[c] (0.94)	−0.27*** (0.00)	−0.29*** (0.00)	0.00***[c] (0.91)	0.04 (0.26)
Paid Help or in Assisted-Living Residence	−0.31*** (0.00)	−0.15* (0.06)	−0.49*** (0.00)	−0.38*** (0.00)	−0.20*** (0.00)	−0.16*** (0.01)
Measures of Patient's Service Use Before Home Health Admission						
Was in Hospital Before Home Health Admission	0.18*** (0.00)	0.20*** (0.00)	0.28*** (0.00)	0.21*** (0.00)	0.20*** (0.00)	0.11** (0.04)
Length of Inpatient Stay Ending in Two Weeks Before Home Health Admission (Days)	0.01** (0.02)	0.00[c] (0.61)	0.01*** (0.00)	0.01** (0.03)	0.00*[c] (0.10)	0.00[c] (0.90)
Whether in SNF in Two Weeks Before Home Health Admission	0.07 (0.18)	0.07 (0.34)	0.17*** (0.00)	0.08 (0.27)	−0.01 (0.90)	−0.06 (0.22)
Home Health Visits from Nondemonstration Agencies in Six Months Before Home Health Admission (Number)	−0.00***[c] (0.00)	−0.00**[c] (0.02)	−0.00*[c] (0.09)	0.00***[c] (0.00)	−0.00***[c] (0.00)	−0.00***[c] (0.00)
Hospitalizations in Six Months Before Home Health Admission (Number)	0.01 (0.71)	0.02 (0.42)	−0.03 (0.12)	−0.07** (0.02)	−0.07*** (0.00)	−0.09*** (0.00)
Time of Home Health Admission						
Quarter Admitted to Home Health Care						
Third quarter 1996	−0.09 (0.18)	0.03 (0.73)	−0.09 (0.22)	0.01 (0.89)	0.00[c] (0.96)	0.12 (0.13)
Fourth quarter 1996	−0.23*** (0.00)	0.13 (0.19)	−0.16** (0.04)	−0.03 (0.80)	−0.08 (0.32)	0.06 (0.44)
First quarter 1997	−0.25** (0.02)	0.17 (0.22)	−0.19 (0.07)	0.17 (0.24)	−0.08 (0.50)	0.18 (0.16)
Second quarter 1997	−0.31*** (0.01)	0.24* (0.06)	−0.20* (0.08)	0.10 (0.48)	−0.16 (0.15)	0.23* (0.06)
Third quarter 1997	−0.30*** (0.01)	0.16 (0.22)	−0.35*** (0.00)	0.13 (0.43)	−0.10 (0.38)	0.25* (0.06)

TABLE C.3 *(continued)*

	Transferring		Oral Medication Management		Pain	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Admitted in Agency's Second or Third Demonstration Year	0.08 (0.49)	-0.20** (0.04)	-0.03 (0.79)	-0.15 (0.29)	0.03 (0.79)	-0.20* (0.07)
Interaction Term of Treatment Status Times Admitted in Agency's Second or Third Demonstration Year	0.11 (0.26)	0.10 (0.41)	0.15 (0.21)	-0.14 (0.25)	0.10 (0.25)	0.00[c] (0.98)
Agency Characteristics						
For Profit	-0.01 (0.95)	-0.02 (0.90)	-0.05 (0.75)	0.09 (0.46)	0.08 (0.49)	0.09 (0.29)
Hospital Based	-0.17 (0.28)	0.14 (0.28)	-0.27* (0.06)	0.03 (0.74)	-0.13 (0.20)	0.21*** (0.01)
Member of a Chain	-0.26*** (0.01)	-0.12 (0.24)	-0.09 (0.36)	-0.08 (0.40)	-0.18** (0.02)	-0.07 (0.26)
Provided Fewer than 30,000 Visits in Base Year	-0.01 (0.93)	-0.10 (0.40)	0.04 (0.79)	0.00[c] (0.99)	-0.24*** (0.01)	-0.03 (0.71)
Predemonstration Practice Pattern (Ratio)[f]	-0.14 (0.46)	0.00[c] (1.00)	0.10 (0.65)	0.10 (0.62)	-0.05 (0.78)	0.18 (0.19)
Area Characteristics						
State						
Florida	-0.16 (0.18)	0.25** (0.03)	-0.06 (0.69)	0.29** (0.01)	-0.01 (0.92)	0.22*** (0.00)
Illinois	-0.14 (0.34)	0.00[c] (0.98)	0.09 (0.54)	0.09 (0.57)	0.13 (0.34)	0.07 (0.50)
Massachusetts	0.33** (0.02)	-0.07 (0.70)	-0.03 (0.86)	0.01 (0.93)	0.12 (0.36)	0.02 (0.88)
Texas	0.04 (0.79)	-0.21 (0.13)	0.13 (0.32)	-0.12 (0.34)	0.19* (0.09)	-0.04 (0.61)

TABLE C.3 *(continued)*

	Transferring		Oral Medication Management		Pain	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Physicians per 10,000 Residents, 1994 (Number)	−0.01 (0.20)	−0.01 (0.14)	0.00[c] (0.62)	0.00[c] (0.30)	0.00[c] (0.39)	0.00[c] (0.34)
Nursing Home Beds per 100 Residents Older than Age 65, 1991 (Number)	−0.02 (0.35)	0.05* (0.08)	−0.03 (0.20)	0.02 (0.26)	−0.05*** (0.00)	−0.01 (0.75)
Hospital Occupancy Rate, 1993	0.86* (0.10)	0.14 (0.82)	0.73 (0.17)	0.48 (0.34)	0.91** (0.02)	0.36 (0.22)
Mean Medicare Reimbursement per Beneficiary, 1991 (Dollars)	0.00[c] (0.28)	0.00[c] (0.16)	0.00[c] (0.11)	0.00[c] (0.42)	0.00**[c] (0.04)	0.00[c] (0.52)
Intercept	0.30 (0.55)	1.14** (0.04)	0.01 (0.99)	1.4*** (0.00)	0.35 (0.41)	1.35*** (0.00)

SOURCE: Quality assurance data.

NOTE: Observations have been weighted to represent agencies equally, and *p*-values have been adjusted for the effects of weighting and clustering. Patients with a previous demonstration home health admission within the year preceding the index admission have been excluded.

[a]Has any of the following risk factors: heavy smoking, obesity, alcoholism, or drug dependency.

[b]By the patient, or by relatives, friend, neighbors, or paid helpers of the patient.

[c]Absolute value of coefficient less than 0.005.

[d]An index of the case-mix-adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

SNF = skilled nursing facility.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

TABLE C.4

ESTIMATED COEFFICIENTS FROM LOGISTIC REGRESSION MODELS OF SELECTED OUTCOMES IN QUALITY ASSURANCE DATA

	Dyspnea		Confusion	
	Improvement (p-Value)	Stabilization (p-Value)	Improvement (p-Value)	Stabilization (p-Value)
Treatment/Control Status				
Agency Was Prospectively Paid	0.06 (0.57)	0.12 (0.18)	0.25** (0.05)	0.17 (0.12)
Demographic Measures				
Original Reason for Medicare: Reached Age 65	0.07 (0.25)	-0.04 (0.61)	0.12 (0.13)	0.03 (0.69)
Age (Years)				
Younger than 65	0.04 (0.57)	0.06 (0.53)	0.18 (0.20)	0.27* (0.06)
75 to 84	-0.13*** (0.00)	-0.09** (0.02)	-0.20*** (0.01)	-0.27*** (0.00)
85 or older	-0.23*** (0.00)	-0.18*** (0.00)	-0.29*** (0.00)	-0.55*** (0.00)
White	-0.03 (0.50)	-0.13** (0.03)	0.03 (0.72)	0.08 (0.37)
Female	0.01 (0.76)	0.03 (0.39)	-0.06 (0.28)	0.08* (0.06)
Has Medicaid Buy-In for Part A and B Medicare	-0.13*** (0.01)	-0.01 (0.87)	-0.00[a] (0.95)	-0.11* (0.10)
Medical Conditions, Symptoms, and Needs at Home Health Admission				
Cancer	-0.13** (0.02)	-0.26*** (0.00)	0.20** (0.02)	0.12* (0.06)
Diabetes	-0.12*** (0.00)	-0.11** (0.02)	0.06 (0.45)	-0.06 (0.25)
Cerebrovascular Accident	0.00[a] (0.99)	-0.05 (0.32)	-0.02 (0.68)	0.16*** (0.01)

TABLE C.4 *(continued)*

	Dyspnea		Confusion	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Need for Complicated Wound Care	0.05 (0.38)	−0.01 (0.93)	0.08 (0.35)	0.05 (0.65)
Functional Limitations				
Eating	−0.03 (0.60)	−0.01 (0.87)	−0.21*** (0.00)	−0.19*** (0.00)
Dressing	0.15*** (0.00)	0.02 (0.60)	−0.23*** (0.00)	−0.14*** (0.01)
Has Risk Factors[b]	−0.05 (0.28)	0.03 (0.57)	−0.01 (0.84)	0.08 (0.22)
Unknown Risk Factors[c]	0.03 (0.68)	−0.07 (0.23)	0.09 (0.30)	−0.07 (0.41)
Medically Unstable	−0.06 (0.11)	−0.12*** (0.00)	0.02 (0.68)	−0.07 (0.27)
Depressed Feelings	0.07 (0.12)	0.07 (0.19)	0.11 (0.17)	−0.02 (0.79)
Displays Depressive Behaviors	0.05 (0.34)	0.01 (0.81)	−0.09 (0.17)	−0.05 (0.45)
Demonstrates Disruptive Behaviors	0.05 (0.33)	0.07 (0.22)	−0.48*** (0.00)	−0.39*** (0.00)
Prognosis at Home Health Admission				
Likelihood that Treatment Can Be Taken Over[d]	0.12*** (0.01)	0.07* (0.10)	0.14** (0.02)	0.12** (0.02)
Prognosis Is Good/Fair	0.09 (0.13)	0.01 (0.85)	0.14* (0.10)	0.07 (0.57)
Life Expectancy Less than Six Months	−0.13* (0.04)	−0.12 (0.11)	−0.00[a] (0.97)	0.14 (0.17)
Rehabilitative Prognosis Is Good	0.16*** (0.00)	0.17 (0.00)	0.15** (0.02)	0.19*** (0.00)

TABLE C.4 *(continued)*

	Dyspnea		Confusion	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Availability of Informal Care at Home Health Admission				
Live-In Informal Help	0.02 (0.56)	−0.05 (0.23)	−0.00* (0.94)	−0.05 (0.24)
Paid Help or in Assisted-Living Residence	−0.10* (0.06)	−0.01 (0.83)	−0.27*** (0.00)	−0.15* (0.06)
Measures of Patient's Service Use Before Home Health Admission				
Was in Hospital Before Home Health Admission	0.10*** (0.01)	0.11*** (0.01)	0.17*** (0.00)	0.18*** (0.00)
Length of Inpatient Stay Ending in Two Weeks Before Home Health Admission (Days)	0.00* (0.32)	0.00* (0.83)	0.00* (0.34)	0.00* (0.61)
Whether in SNF in Two Weeks Before Home Health Admission	0.11** (0.02)	0.08* (0.08)	0.23*** (0.00)	0.05 (0.44)
Home Health Visits from Nondemonstration Agencies Six Months Before Home Health Admission (Number)	−0.00* (0.23)	−0.00** (0.04)	−0.00** (0.09)	−0.00** (0.07)
Hospitalizations in Six Months Home Health Admission (Number)	−0.01 (0.54)	−0.02 (0.28)	0.02 (0.50)	−0.04 (0.13)
Time of Home Health Admission				
Quarter Admitted to Home Health Care				
Third quarter 1996	−0.04 (0.59)	0.02 (0.70)	−0.06 (0.46)	−0.03 (0.71)
Fourth quarter 1996	−0.09 (0.21)	0.05 (0.50)	−0.01 (0.90)	0.05 (0.63)
First quarter 1997	−0.16 (0.20)	0.17* (0.10)	−0.12 (0.38)	0.07 (0.62)
Second quarter 1997	−0.19 (0.10)	0.25** (0.020	−0.16 (0.23)	0.13 (0.37)
Third quarter 1997	−0.20 (0.11)	0.24** (0.02)	−0.15 (0.20)	0.16 (0.26)

TABLE C.4 *(continued)*

	Dyspnea		Confusion	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Admitted in Agency's Second or Third Demonstration Year	0.08 (0.45)	-0.18* (0.06)	0.03 (0.78)	-0.17 (0.17)
Interaction Term of Treatment Status Times Admitted in Agency's Second or Third Demonstration Year	0.09 (0.27)	-0.01 (0.91)	-0.01 (0.54)	-0.05 (0.63)
Agency Characteristics				
For Profit	-0.10 (0.50)	0.03 (0.78)	-0.12 (0.47)	0.03 (0.86)
Hospital Based	-0.23 (0.11)	0.10 (0.38)	-0.14 (0.33)	0.27* (0.08)
Member of a Chain	-0.02 (0.80)	-0.13 (0.15)	-0.08 (0.51)	-0.15 (0.19)
Provided Fewer than 30,000 Visits in Base Year	-0.22 (0.11)	-0.12 (0.28)	0.03 (0.85)	0.17 (0.24)
Predemonstration Practice Pattern (ratio)†	-0.13 (0.35)	0.18 (0.28)	0.05 (0.84)	0.18 (0.50)
Area Characteristics				
State				
Florida	-0.20 (0.14)	0.12 (0.25)	-0.11 (0.52)	0.18 (0.18)
Illinois	0.26 (0.11)	0.09 (0.49)	0.13 (0.43)	0.06 (0.71)
Massachusetts	0.11 (0.50)	-0.12 (0.48)	-0.20 (0.18)	-0.27 (0.16)
Texas	0.21* (0.07)	-0.13 (0.31)	0.18 (0.22)	-0.13 (0.41)
Physicians per 10,000 Residents, 1994 (Number)	0.01 (0.20)	0.00* (0.77)	0.02** (0.04)	0.00* (0.52)
Nursing Home Beds per 100 Residents Older than Age 65, 1991 (Number)	-0.08*** (0.00)	0.01 (0.83)	-0.11*** (0.00)	0.00* (0.95)

TABLE C.4 *(continued)*

	Dyspnea		Confusion	
	Improvement (*p*-Value)	Stabilization (*p*-Value)	Improvement (*p*-Value)	Stabilization (*p*-Value)
Hospital Occupancy Rate, 1993	0.37 (0.42)	0.15 (0.75)	0.54 (0.36)	0.49 (0.21)
Mean Medicare Reimbursement per Beneficiary, 1991 (Dollars)	-0.00*** (0.02)	-0.00* (0.36)	-0.00** (0.08)	-0.00* (0.59)
Intercept	0.77 (0.15)	1.49*** (0.00)	0.42 (0.51)	1.63*** (0.01)

SOURCE: Quality assurance data.

NOTE: Observations have been weighted to represent agencies equally, and *p*-values have been adjusted for the effects of weighting and clustering. Patients with a previous demonstration home health admission within the year before the index admission have been excluded.

[a]Absolute value of coefficient less than 0.005.

[b]Has any of the following risk factors: heavy smoking, obesity, alcoholism, or drug dependency.

[c]A zero or one indicator variable with the value one is risk factor information was missing. If risk factor information was missing, an imputed value was used for "has risk factors".

[d]By the patient, or by relatives, friend, neighbors, or paid helpers of the patient.

[e]An index of the case-mix-adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

SNF = skilled nursing facility.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.

TABLE C.5

ESTIMATED COEFFICIENTS FROM LOGISTIC REGRESSION MODELS OF SELECTED SATISFACTION OUTCOMES IN FOUR MONTH SURVEY

	Staff Rushed Through Work Most or All of the Time (*p*-Value)	Staff Paid Attention to Patient Some of the Time, or Little or None of the Time (*p*-Value)	Staff Encouraged Independence--Disagree or Strongly Disagree (*p*-Value)
Treatment/Control Status			
Agency Was Prospectively Paid	0.47**	0.61***	0.44**
	(0.03)	(0.01)	(0.04)
Demographic Measures			
Original Reason for Medicare: Reached Age 65	-0.36	-0.74**	0.29
	(0.36)	(0.04)	(0.50)
Age (Years)			
Younger than 65	0.37	-0.05	0.62
	(0.50)	(0.90)	(0.23)
75 to 84	0.56*	0.30	0.41
	(0.06)	(0.31)	(0.11)
85 or older	0.64**	0.47	0.30
	(0.05)	(0.19)	(0.29)
White	-0.73**	0.25	-0.69**
	(0.02)	(0.54)	(0.02)
Female	-0.15	-0.10	0.14
	(0.62)	(0.72)	(0.60)
Has Medicaid Buy-In for Part A and B Medicare	0.42	0.19	0.08
	(0.16)	(0.53)	(0.72)
Either in HMO or Medicare Secondary Payer at Some Time in Six Months Before Home Health Admission	0.79*	-0.40	-0.23*
	(0.10)	(0.49)	(0.70)
Patient Was Survey Respondent	0.47	0.20	0.37*
	(0.12)	(0.47)	(0.09)
Medical Conditions, Symptoms, and Needs at Home Health Admission			
Cancer	0.23	0.32	0.56*
	(0.50)	(0.35)	(0.06)

TABLE C.5 *(continued)*

	Staff Rushed Through Work Most or All of the Time (*p*-Value)	Staff Paid Attention to Patient Some of the Time, or Little or None of the Time (*p*-Value)	Staff Encouraged Independence--Disagree or Strongly Disagree (*p*-Value)
Diabetes	0.08 (0.75)	-0.26 (0.35)	-0.31 (0.23)
Cerebrovascular Accident	0.03 (0.94)	0.73** (0.05)	0.62*** (0.01)
Stage 3 or 4 Decubitus Ulcers	-0.97 (0.27)	-0.23 (0.67)	-1.23* (0.08)
Need for Complicated Wound Care	0.29 (0.61)	1.30*** (0.00)	0.22 (0.64)
Functional Limitations			
Bathing	-0.20 (0.65)	-0.11 (0.78)	0.22 (0.48)
Eating	-0.46 (0.19)	0.37 (0.26)	0.62** (0.04)
Dressing	-0.25 (0.51)	0.26 (0.53)	-0.28 (0.32)
Toileting	0.72* (0.08)	-0.08 (0.83)	0.41 (0.32)
Transferring	0.13 (0.74)	-0.12 (0.73)	-0.50 (0.19)
Availability of Informal Care at Home Health Admission			
Was Married at Time of Home Health Admission	-0.04 (0.90)	-0.02 (0.94)	-0.11 (0.61)
Had Non-Medicare Homemaker or Aide in Month Before Home Health Admission	0.08 (0.84)	0.20 (0.53)	0.26 (0.38)
Had (Nonpaid) Help from Friends and Family in Month Before Home Health Admission	0.11 (0.77)	0.40 (0.17)	0.36 (0.15)
Had (Nonpaid) Friend of Family Living at Home in Month Before Home Health Admission	-0.03 (0.94)	0.19 (0.45)	-0.19 (0.48)

TABLE C.5 *(continued)*

	Staff Rushed Through Work Most or All of the Time (*p*-Value)	Staff Paid Attention to Patient Some of the Time, or Little or None of the Time (*p*-Value)	Staff Encouraged Independence--Disagree or Strongly Disagree (*p*-Value)
Measures of Patient's Service Use Before Home Health Admission			
Was in Hospital Before Home Health Admission	0.72*** (0.01)	0.11 (0.72)	0.03 (0.90)
Length of Inpatient Stay Ending in Two Weeks Before Home Health Admission (Days)	-0.05* (0.10)	0.01 (0.56)	-0.00[a] (0.94)
Whether in SNF in Two Weeks Before Home Health Admission	-0.13 (0.70)	-0.69* (0.09)	-0.33 (0.35)
Home Health Visits from Nondemonstration Agencies in Six Months Before Home Health Admission (Number)	0.01** (0.05)	0.01** (0.04)	0.00[a] (0.31)
Whether Hospitalized in Six Months Before Home Health Admission	0.29 (0.39)	0.20 (0.59)	-0.16 (0.56)
Total Medicare Part A Reimbursement Before Home Health Admission (in 1000s of Dollars)	-0.01 (0.42)	-0.01 (0.64)	0.00[a] (1.00)
Quarter Admitted to Home Health Care			
Second Quarter 1997	0.24 (0.28)	-0.32 (0.17)	0.32 (0.16)
Third Quarter 1997	-0.46 (0.24)	-0.85*** (0.01)	-0.08 (0.76)
Agency Characteristics			
For Profit	0.41 (0.15)	-0.84*** (0.00)	-0.36 (0.24)
Hospital Based	0.07 (0.79)	0.23 (0.38)	0.05 (0.87)
Member of a Chain	-0.59* (0.06)	0.32 (0.24)	0.51*** (0.03)

TABLE C.5 *(continued)*

	Staff Rushed Through Work Most or All of the Time (*p*-Value)	Staff Paid Attention to Patient Some of the Time, or Little or None of the Time (*p*-Value)	Staff Encouraged Independence--Disagree or Strongly Disagree (*p*-Value)
Provided Fewer than 30,000 Visits in Base Year	-0.79***	0.18	-0.23
	(0.00)	(0.48)	(0.46)
Predemonstration Practice Pattern (Ratio)[†]	-0.41	0.75*	-0.01
	(0.36)	(0.08)	(0.99)
Area Characteristics			
State			
Florida	-0.67	-0.32	-0.25
	(0.14)	(0.39)	(0.43)
Illinois	-0.01	0.68*	0.46
	(0.99)	(0.06)	(0.13)
Massachusetts	-0.44	-1.17**	-0.10
	(0.37)	(0.02)	(0.84)
Texas	0.53*	0.07	0.43
	(0.09)	(0.82)	(0.14)
Urban Area	-0.24	0.65*	-0.22
	(0.58)	(0.08)	(0.61)
Physicians per 10,000 Residents, 1994 (Number)	0.05***	0.00[a]	0.00[a]
	(0.01)	(0.97)	(0.90)
Nursing Home Beds per 100 Residents Older than Age 65, 1991 (Number)	-0.17**	0.04	-0.06
	(0.04)	(0.57)	(0.33)
Hospital Occupancy Rate, 1993	-0.47	1.02	0.01
	(0.72)	(0.49)	(0.99)
Mean Medicare Reimbursement per Beneficiary, 1991 (Dollars)	-0.00***	-0.00[a]	0.00[a]
	(0.01)	(0.98)	(0.72)
Intercept	0.36	-5.6***	-3.30**
	(0.84)	(0.00)	(0.04)

SOURCE: Four-month patient survey.

NOTE: Observations have been weighted to represent agencies equally, and *p*-values have been adjusted for the effects of weighting and clustering. Patients with a previous demonstration home health admission within the year before the index solution have been excluded.

TABLE C.5 *(continued)*

[a] Value of coefficient less than 0.005.

[b] An index of the case-mix-adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

*Significantly different from zero at the 0.10 level, two-tailed test.
**Significantly different from zero at the 0.05 level, two-tailed test.
***Significantly different from zero at the 0.01 level, two-tailed test.